COMPARABILITY AND TRANSFERABILITY IN ECOSYSTEM-ASSESSMENT TECHNIQUES AND TOOLS: AN INTERNATIONAL CASE STUDY

by

Kyung Seo Park

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Natural Resources and Environment) in the University of Michigan 2015

Doctoral Committee:

Professor Michael J. Wiley, Co-Chair
Professor James S. Diana, Co-Chair
Professor J. David Allan
Professor Kwang-Guk An, Chungnam National University
Professor Aline Cotel

To Jessica and Kun-Yong,

My family is always my number one priority.

And to Brothers, Mom and Dad;

Your endless support, encouragement, sacrifice, and love
allowed me to harvest crops today.

Acknowledgements

I would like to express special thanks to my advisors and co-chairs Drs. Michael Wiley and James Diana. They have been great and patient mentors, guides, supporters and editors of my work. They have taught me not only how to fish, but also why I fish and how beneficially I can use the fish. I would also like to thank the wives of my advisors, Susan Wiley and Barbara Diana, for their kindness, love and help to enjoy my graduate life in Ann Arbor with joy and hapiness.

Many people contributed to this dissertation. This dissertation would not have been written without my committee members. Special thanks to Drs. Dave Allan, Aline Cotel, and Kwang-Guk An for their helpful feedback and broader perspective on my research. I would like to acknowledge Drs. Catherine Riseng, Sara Adlerstein, and Rachel Kaplan, who helped and guided me with their invaluable research experience and humors to make my graduate life successful and enjoyable.

I would like to thank my previous Korean advisors and mentors, who led and inspired me to study abroad with a big dream and hope since my undergraduate life. Special thanks to Shin-Suk Choi, Young-Pyo Hong, Sung-Min Boo, Seong-Deog Kim, Kwang-Guk An, Dong-Hyuk Yeom, Jong-Hoon Shim, Jae-Gu Kim, Sung-Ja Bang, Min-ho Jang, Doo-shin Park and Jin-Won Seo. Withouth their help and continuous support, I couldn’t dream high and achieve my goal.

I have been very fortunate with a diverse and talented group of fellow graduate students and researchers, who shared their experience and paid attention to my complaints and questions. I would like to acknowledge Shaw Lacy, Ethan Bright, Beth Spark-Jackson, Glenn Carter, Solomon David, Yu-Chun Kao, Suting Cheng, Michael Fainter, Andy James, Lauren Eaton, Lori Ivan, Min-Wook Kim, Ji Woong Choi, Jae Hoon Lee, Woon-Ki Moon and field sampling crews in Michigan and Korea

I am also indebted to many SNRE staffs. Especially, I would like to acknowledge Diana Woodworth, Jennifer Taylor, Judy Byington, Lisa Yee-Litzenberg, Sucila Fernandes, Jeff Keeler, Sondra Auerbach. They always provided me the best environment to focus on my research by supporting my academic, financial and legal process without hesitation.

Lastly, I am deeply thankful to my loving family and parents, who stood by me, assisted me, and encouraged me all the time. Heartful and respectful thanks to my wife (Kun-Yong Kim), daughter (Jessica Park), brothers (Chan-Seo Park and Jong-Seo Park) and parents (Geum Ja Lee and Seung Il Park) for making me moving forward.

Table of Contents

Dedication .... **ii**

Acknowledgements .... **iii**

List of Tables .... **x**

List of Figures .... **xv**

Abstract .... **xx**

Chapter 1 : Disseratation introduction and overview **1**

Overview of ecosystem-assessment techniques and tools 1

Global transferability of EATTs 2

The ecological setting of Michigan and S. Korean rivers 6

Research goals and questions 8

Dissertation content 9

Literature cited 18

Chapter 2 : Estimation and classification of flow regimes for South Korean streams and rivers **23**

Abstract 23

Introduction 25

Materials and methods 28

Characteristics of Korean streams and rivers 28

Data collection and summary 28

Multiple linear regression model development 30

Performance evaluation of four different predictive models 31
Korean stream flow-type classification methods 32
Statistical analysis 33

Results 35
MLR models 35
Evaluation of model performance 36
Classification of Korean stream flow types 38
Stream flow-type classification of nation-wide biological sampling sites in S. Korea .. 39

Discussion 41
Factors controlling stream flows in S. Korea 42
Classification of Korean stream flow types 43
Assumptions and limitations 44
Management implications 46

Literature cited 73

Chapter 3 : The predictive performance evaluation of four different analytical methods in the estimation of Korean stream flow regimes from landscape variables 80

Abstract 80

Introduction 81

Materials and methods 84
Data collection and summary 84
Modeling approach: MLR, PCR, ANN, and PC-ANN 86
Multiple linear regressions (MLR) 87
Principal component regressions (PCR) 87
Artificial neural network (ANN) 87
Principal component-artificial neural network (PC-ANN) 88
Performance evaluation of four different predictive models 88

Results 91
Summary statistics of response and explanatory variables 91
Predictive models for 10% exceedance discharge (high flows) 92
Predictive models for 50% exceedance discharge (median flows) 93

Predictive models for 90% exceedance discharge (low flows) 93
Discussion 95
Linear regression models (MLR and PCR) 96
Non-linear models (ANN and PC-ANN) 96
Error in databases and other limitations 97
Literature cited 110

Chapter 4 : Biases arising from sampling gear in stream bioassessments: electrofishing versus cast-netting 116

Abstract 116
Introduction 118
Materials and methods 121
Study site and periods 121
Sampling distance and time 122
Electrofishing 122
Cast-netting 123
Sampling crews and fish identification 123
Fish indicator metrics and data analysis 124
Statistical analysis 125
Results 126
Characteristics of stream sites and fish habitats 126
Comparison of sampled fish taxa richness, abundance, and occurrence 126
Frequency of occurrence 128
Comparison of fish sampling gear efficiency (FSGE) 129
Impacts of environmental factors 130
Discussion 132
Ecological significances and implications 135
Literature cited 155

Chapter 5 : Biases arising from benthic macroinvertebrate methods in rapid bioassessment programs (RBPs) for wadeable streams: Michigan versus Korean assessment protocols 158

Abstract 158

Introduction 159

Materials and methods 163

Study sites and periods 163

Benthic macroinvertebrate sampling 164

Michigan department of environmental quality procedure 51 (MDEQP51) sampling 165

Korean nationwide aquatic ecological monitoring program (KNAEMP) sampling 165

Identification of benthic macroinvertebrates 166

Benthic macroinvertebrate metrics and data analysis 166

Statistical analysis 168

Results 169

Characteristics of stream sites and benthic macroinvertebrate habitats 169

Comparison of sampled taxa richness, abundance, and occurrence 169

Comparison of macroinvertebrate sampling method efficiency (SME) 171

Metric comparability 172

Discussion 175

Assessing impacts with biased indicators 178

Summary and conclusion 179

Literature cited 196

Chapter 6 : Meta-analysis with ecological assessment data: a case study of Michigan and South Korean streams 202

Abstract 202

Introduction 203

Materials and methods 207

Case study approach 207

Fish sampling methods 208

Benthic macroinvertebrate sampling methods 209

Indicator metrics 210

Environmental data 212

Region-specific ecological normalization of assessment metrics 214

Statistical analysis 216

Results 217

Differences in ecological context 217

LU stressor-response relationships: individual indicator metrics 219

LU stressor-response relationships: multi-metrics 220

Comparison of ecological condition in Michigan and Korean streams 222

Discussion 224

Sampling method biases 224

Correcting for ecological covariance 225

LU stressor-response relationships 228

Implications for global method transferability and inter-regional assessment 229

Summary 231

Literature cited 266

List of Tables

Table 1.1. Summary of general information of Michigan and S. Korea. Population densities for Michigan and S. Korea were summarized with 2010 summary data obtained from US Census Bureau 2015 and KOSIS 2015. 13

Table 1.2. Percent taxonomic overlap between the two regions. Genera and Species for benthic macroinvertebrates were not summarized here because MDEQ data were collected at the family level of identification. Taxa list was based on the data of National Institute for Environmental Research, S. Korea (NIER 2009) and Michigan Department of Environmental Quality, USA (MDEQ 1997). 14

Table 2.1. Summary of watershed characteristics of the five major river watersheds and site numbers used for the estimation and classification for stream flow regimes of stream flow gages and un-gauged biological sampling sites used by the National Aquatic Ecological Monitoring Program (NAEMP) throughout South Korea. 49

Table 2.2. Summary statistics of stream discharges (cms) at each percent exceedance frequency in Korean streams and rivers used for normalizing linear regression models as dependent variables. Stream discharge is represented by "Q" and defined as 05-95% exceedance discharges (cms). "n" indicates the total number of gages for the summary of stream discharges and "SD" indicates standard deviation. 50

Table 2.3. Summary statistics of candidate independent variables used for multiple linear regression models of a series of exceedance frequencies for annual and four seasonal time windows. 51

Table 2.4. Summary statistics of 5 and 95 percent exceedance discharges for each major watershed. Stream discharge is represented by "Q" and "n" indicates the total number of gages for the summary of stream discharges and "SD" indicates standard deviation. 52

Table 2.5. Multiple linear regression models of all exceedance flows for annual and seasonal time periods. Bold indicates significance at $p \leq 0.05$, and bold and italics indicate significance at $p \leq 0.01$ 53

Table 2.6. Table 6. Mean absolute percentage errors (MAEs) and Nash-Sutcliffe model efficiency coefficient (NSE) of exceedance flow models for annual and four seasonal time windows. 55

Table 2.7. Two tailed Pearson correlation tests between observed and predicted stream discharges for all sites combined (n=163) and four major watersheds. Bold indicates significance at $p \leq 0.05$, and bold and italics indicate significance at $p \leq 0.01$. 56

Table 2.8. Principal components analysis (PCA). Stream flow characteristics variable loadings for PC 1, 2, and 3 (n= 163). Bold values are considered equal or larger than |0.40|, but PC3 was not considered because its eigen value (0.190) was less than 1 57

Table 2.9. Comparison of proportion (site numbers) of stream flow types by flow stability and stream power for stream flow gages and biological sampling sites. 58

Table 2.10. Comparison of proportion (site numbers) of each stream flow type for stream flow gages and biological sampling sites. 59

Table 3.1. Summary statistics of dependent and independent variables used for comparison of predictive models. All other variables not used for model construction were removed from the table. SD indicates standard deviation. 100

Table 3.2. Mean absolute percentage errors (MAEs) and Nash-Sutcliffe model efficiency coefficient (NSE) of each model for three percent exceedance discharges (Q10, Q50, and Q90). 101

Table 3.3. Two tailed Pearson correlation tests among observed and predicted stream discharges for all sites combined (n=163). Bold indicates significance at $p \leq 0.05$, and bold and italics indicate significance at $p \leq 0.01$. 102

Table 3.4. Regression coefficients and collinearity statistics of the MLR and PCR models for 10%, 50%, and 90% exceedance discharges (Q10, Q50, and Q90). Bold unstandardized coefficients indicate significance at $p \leq 0.05$, and bold and italics unstandardized coefficients indicate significance at $p \leq 0.01$. 103

Table 3.5. Collinearity statistics for MLR and PCR models. Variance inflation factors (VIFs) and average VIFs were summarized in order to see multicollinearity issues among independent variables used in MLR and PCR models. 104

Table 4.1. Summary statistics of environmental factors that may have significant influences on fish sampling performance with electrofishing and cast-netting sampling gears. SD indicates standard deviation. 136

Table 4.2. Fish indicator metrics used for the Korean Index of Biological Integrity using Fish (K-IBIF), the South Korean National River Assessment Program (NIER 2009), the biological monitoring program of the Michigan Department of Environmental Quality (MDEQ Procedure-51) (MDEQ 2002), and this study 137

Table 4.3. Fish species composition, individual numbers collected, and site occurrence frequency of each species for cast-netting (CN), electrofishing (EF), and all data from both gear (AD)... 138

Table 4.4. Summary statistics of fish species for all fish indicator metrics used for fish species data comparability (n = 21)... 140

Table 4.5. Summary statistics of paired samples T-tests among three data sets collected by different sampling gears. CI and df indicate confidence interval and degree of freedom, respectively. *=0.05; **=0.001.. 141

Table 4.6. Two tailed Pearson correlation tests between fish sampling data from different gears for each fish metric. Three fish groups removed were Centrachidae, Esociformes, and Petromyzontiformes. Bold indicates significance at p≤0.05, and bold and italics indicate significance at p≤0.01. “a” indicates no enough data to compare... 143

Table 4.7. Comparison of fish species occurrence ratios among cast-netting (CN), electrofishng (EF), and all data from both gear (AD). The fish species that occurred in less than 4 sites in all data from both gear (AD) were removed from the table and the occurrence ratios were sorted from largest to smallest.. 144

Table 4.8. P values from ANCOVA tests of SGEs for fish metrics. Sampling gear was used as fixed factor and urban and agricultural land uses were used as a covariate. 145

Table 4.9. Summary statistics of paired samples T-tests among three data sets collected by different sampling gears. Fish species of Esociforms, Centrachidae, and Petromyzontidae were excluded in this analysis. CN, EF, and AD indicate cast-netting, electrofishing, and combined data, respectively. CI and df indicate confidence interval and degree of freedom, respectively. *=0.05; **=0.001.. 146

Table 5.1. Summary statistics of stream habitat characteristics and landuse patterns for benthic macroinvertebrate sampling sites. SD indicates standard deviation... 181

Table 5.2. Macroinvertebrate indicator metrics and their abbreviations used for this study. * indicates that metrics were used for the evaluation of sampling gear efficiency. 182

Table 5.3. Comparison of benthic macroinvertebrate data for specific orders. Data were summarized by MDEQP51, KAEMAP, and COMB (combined data of MDEQP51 and KAEMAP). 183

Table 5.4. Summary statistics of all macroinvertebrate indicator metrics used for macroinvertebrate data comparability (n = 29).. 184

Table 5.5. Two tailed Pearson correlation tests between KNAEMP and Michigan MDEQP51. Bold indicates significance at p≤0.05, and bold and italics indicate significance at p≤0.01..... 185

Table 5.6. Summary statistics of paired samples T-tests (KNAEMP versus MDEQP51). CI and df indicate confidence interval and degree of freedom, respectively. *=0.05; **=0.001........... 186

Table 5.7. P values from ANCOVA tests for benthic macroinvertebrate metrics. Each sampling method indicated KNAEMP and MDEQP51 and was used as fixed factor. Also, urban and agricultural land uses were used as covariates to observe biological response to land uses. 187

Table 6.1. Definition of acronyms frequently used in this study. 233

Table 6.2. List of biological indicator metrics (fish and benthic macroinvertebrates) used for this study and their equations employed to each metric to correct numerical differences by sampling gear or methods between Michigan and S. Korean regions. MMI and MKO indicate individual metrics of Michigan and S. Korea, respectively. cMMI and cMKO indicate method corrected individual metrics of each region. 234

Table 6.3. Summary statistics of landscape variables between Michigan and South Korean regions used for the study. SD indicates standard deviation. 235

Table 6.4. Pearson correlation coefficients between biological indicator metrics and LU stressors for Michigan, Korean (large river sites excluded), and combined (both regions) datasets. Bold indicates significance at $p \leq 0.05$ and bold and italics indicate significance at $p \leq 0.01$. 236

Table 6.5. Coefficients and their statistical significance from GLM ANCOVA tests of normalized scores of biological indicator metrics with urban and agricultural LUs between Michigan and Korean (large river site excluded) datasets. Region (Michigan and Korea) was used as fixed factor and LU stressors (xUrb and xAg) were used as covariates. Bold indicates significance at $p \leq 0.05$ and bold and italics indicate significance at $p \leq 0.01$. Coefficients for Region, Region*ln(xUrb), and Region*ln(xAg) were summarized for Korean data and coefficients of these for Michigan data could be calculated by multiplying by -1 to coefficients of Korean data. 237

Table 6.6. Pearson correlation coefficients between assessment metrics (raw and normalized) and LU stressors for Michigan, Korean, and combined (both regions) datasets. Large river sites were removed for Korean data. Bold indicates significance at $p \leq 0.05$ and bold and italics indicate significance at $p \leq 0.01$. 238

Table 6.7. Coefficients, estimated slopes, and slope difference calculated from GLM ANCOVA tests of normalized assessment scores with urban and agricultural LUs between Michigan and Korean datasets. Region (Michigan and Korea) was used as fixed factor and LU stressors (xUrb and xAg) were used as covariates. Estimated slopes of were estimated from coefficient and interaction coefficient with region. Slope difference was calculated by estimated slope of Korea divided by estimated slope of Michigan. Statistical significance cut-off was 0.1. Bold coefficient values indicate significance at $p \leq 0.10$ and bold and italic values indicate significance at $p \leq 0.05$.. 239

Table 6.8. Percentages of sites for each stream health class for normalized raw and composite multimetric indices.. 240

Table 6.9. Summary statistics of normalized-raw and -composite multimetrics for Michigan and Korea regions. SD indicates standard deviation. 241

Table 6.10. Percentages of impaired streams sites for small and large streams based on stream health classification with raw and normalized assessment scores. Impairment classification of raw assessment scores for fish and benthic macroinvertebrates were based on P51Fi and P51Inv classes for Michigan region and KIBI and KSI classes for S. Korean region. These regional data were regionally normalized and reclassified for the impairment status based on normalized assessment scores. Large stream sites were arbitrarily defined as having catchment areas bigger than 500 km^2. SM and LR stand for smaller and larger streams, respectively. 242

List of Figures

Figure 1.1. Map showing location of Michigan and Korean regions, which lies in the same 'temperate seasonal forest' biome. The picture is redrawn from Ricklefs 2008. 15

Figure 1.2. Comparison of average monthly air temperature between Lower Peninsula Michigan and S. Korea. Average monthly air temperature data for Michigan and S. Korea were summarized with data from 1981 to 2010, obtained from National Environmental Satellite, Data, and Information Service (NESDIS 2011) and Korea Meteorological Administration (KMA 2011), respectively. 16

Figure 1.3. Comparison of average monthly precipitation between Lower Peninsula Michigan and S. Korea. Average monthly precipitation data for Michigan and S. Korea were summarized with data from 1981 to 2010, obtained from National Environmental Satellite, Data, and Information Service (NESDIS 2011) and Korea Meteorological Administration (KMA 2011), respectively. 17

Figure 2.1. Locations of A) stream discharge gages (163 sites) and B) weather-observation stations (63 sites). Violet, sky blue, orange, and yellow colors indicate the Han River, Geum River, Nakdong River, and Youngsum River (Youngsan and Sumjin Rivers) Watersheds, respectively. 60

Figure 2.2. Comparison of mean stream discharges (cms) of percent exceedance freqencies for five time windows (Annual, January, May, July, and October) for all watersheds. 61

Figure 2.3. Comparison of mean stream discharges (cms) of each percent exceedance frequency for overall and four major Korean watersheds (Geum River, Han River, Nakdong River, and Youngsum (Youngsan and Sumjin) River). 62

Figure 2.4. Scatter plots of the relationship between predicted and observed stream discharges for Annual time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line. 63

Figure 2.5. Scatter plots of the relationship between predicted and observed exceedance discharges for January time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line. 64

Figure 2.6. Scatter plots of the relationship between predicted and observed exceedance discharges for July time window. Left and right column panels show the test group (133 sites)

and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line.. 65

Figure 2.7. Scatter plots of the relationship between predicted and observed exceedance discharges for May time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line.. 66

Figure 2.8. Scatter plots of the relationship between predicted and observed exceedance discharges for October time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line... 67

Figure 2.9. Two-dimensional PC score plot for the Korean stream classification using stream power and flow stability. Median numbers of both PC scores were used for stream classification. F, S, HP, and LP indicate flashier stream, more stable stream, higher-powered stream, and lower-powered stream, respectively.. 68

Figure 2.10. Example flow duration curves of four different stream flow types for Korean streams based on stream classification using PCA analysis. F, S, HP, and LP indicate flashier stream, more stable stream, higher-powered stream, and lower-powered stream, respectively... 69

Figure 2.11. Comparison of landscape variables and stream flow characteristics among four different stream flow types for stream flow gages.. 70

Figure 2.12. Comparison of site numbers for four different flow types between stream flow gages and biological monitoring sites. Stream flow type for each site was produced by PCA and PCR models. F, S, HP, and LP indicate flashier stream, more stable stream, higher power stream, and lower power stream, respectively. 71

Figure 2.13. Comparison of percentages of each flow type for major four Korean watersheds. Top graph shows stream flow gages (n=163) and bottom graph shows biological monitoring sites (n=684 sites). 72

Figure 3.1. Locations of A) stream discharge gages (163 sites) and B) weather-observation stations (63 sites).. 105

Figure 3.2. Boxplots showing the mean (dashed line), median (solid line), and range of stream discharges (cms) for exceedance discharges (Q10, Q50, and Q90) used for predictive models. Y-axis is log scaled for better view of discharge distribution... 106

Figure 3.3. Scatter plots of observed and predicted 10% exceedance flows (Q10). Predicted exceedance flows were modeled by ANN, MLR, PC-ANN, and PCR. 107

Figure 3.4. Scatter plots of observed and predicted 50% exceedance flows (Q50). Predicted exceedance flows were modeled by ANN, MLR, PC-ANN, and PCR. 108

Figure 3.5. Scatter plots of observed and predicted 90% exceedance flows (Q90) Predicted exceedance flows were modeled by ANN, MLR, PC-ANN, and PCR. 109

Figure 4.1. Locations of fish sampling sites (n= 21) for electrofishing (EF) and cast-netting (CN). 148

Figure 4.2. Comparison of species numbers between cast-netting (CN) and electrofishing (EF) and correction equations for sampling gear bias.. 149

Figure 4.3. Comparison of individual number percentages between cast-netting (CN) and electrofishing (EF) and correction equations for sampling gear bias. 150

Figure 4.4. Site occurrence frequencies and relative fish abundances of the highest top ten fish species for cast-netting method (CN), electrofishing method (EF), and all data from both gear (AD). 151

Figure 4.5. Mean fish sampling gear efficiency (SGE) with 95 percentile ranges for cast-netting and electrofishing methods. A) SGE summarized for each order (CYNI: Cypriniformes, ESOC: Esociformes, PERC: Perciformes, PETR: Petromyzontiformes, SCOR: Scorpaeniformes, SILU: Siluriformes). B) SGE summarized for each family (CATO: Catostomidae, CENT: Centrachidae, COTT: Cottidae, CYPR: Cyprinidae, ESOC: Esocidae, ICTA: Ictaluridae, PERC: Percidae, PETR: Petromizontidae). Orders and families collected in less than seven sites were removed from the graphs. 152

Figure 4.6. Mean fish sampling gear efficiency (SGE) with 95 percentile ranges for cast-netting and electrofishing methods. SGE summarized for each fish indicator metric........................... 153

Figure 4.7. Scatter plots showing the relationship between sampling gear efficiency and environmental factors for each sampling gear using total number of species. 154

Figure 5.1. Locations of benthic macroinvertebrate sampling sites (n= 29) in the Geum River Watershed, South Korea. 188

Figure 5.2. Site occurrence frequencies and relative abundances of the top ten macroinvertebrate taxa by KNAEMP, MDEQP51, and COMB (combined data). Left column has site occurrence frequencies of each taxa and right column has relative abundances of each taxa. Order taxa were used for class Insecta and class taxa were used for others.. 189

Figure 5.3. Mean macroinvertebrate sampling methodology efficiency (MSME) with 95 percentile ranges for KNAEMP and MDEQP51. 190

Figure 5.4. Mean macroinvertebrate sampling methodology efficiency (MSME) with 95 percentile ranges for KNAEMP and MDEQP51. 191

Figure 5.5. Comparison of taxa richness of each metric between KNAEMP and MDEQP51 and correction equations for sampling method bias. 192

Figure 5.6. Comparison of benthic macroinvertebrate metric values between KNAEMP and MDEQP51 and correction equations for sampling method bias. 193

Figure 5.7. Comparison of taxa richness of two biological monitoring programs against proportion of urban landuse for each taxonomic resolution. Graphs on the left columns show data set before the standardization of taxa richness and graphs on the right columns show data set after the standardization of taxa richness. 194

Figure 5.8. Comparison of taxa richness of two biological monitoring programs against proportion of agricultural landuse for each taxonomic resolution. Graphs on the left columns show data set before the standardization of taxa richness and graphs on the right columns show data set after the standardization of taxa richness. 195

Figure 6.1. Pie charts of stream health assessment classification of Korean streams (n= 633) produced by each regional rapid bioassessment protocols for fish and benthic macroinvertebrates. All assessment data were obtained from the Korean National Aquatic Ecological Monitoring Program (NIER 2009). 243

Figure 6.2. Comparison of the frequency distribution of site drainage-area (ln(drainage area+1)) for each region (A: Michigan region (803 sites), B: Korean region with large river sites included (684 sites), and C: Korean region with large river sites excluded (633 sites). 244

Figure 6.3. Comparison of the urban and agricultural LU frequency distribution between Michigan (n= 803) and Korean (n=684) regions. 245

Figure 6.4. Comparison of the distribution of landscape variables against site drainage area between Michigan and Korean regions. 246

Figure 6.5. Comparision of high (10%), median (50%), and low (90%) frequency flow yields ($m^3/km^2/sec$) between Michigan and Korean regions. 247

Figure 6.6. Comparison of Michigan (blue circles) and Korean (red x marks) datasets. Axes are A) number of total fish species (nFiSp) and B) number of total invertebrate families (nFaInv) plotted against natural log of site drainage area. All fish and invertebrate data were from raw data collected by regional sampling methods. Large river sites for Korean dataset were excluded. . 248

Figure 6.7. Scatter plots of number of total fish species (nFiSp) against urban and agricultural LUs between Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data). 249

Figure 6.8. Scatter plots of number of total invertebrate families (nFaInv) against urban and agricultural LUs Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data). 250

Figure 6.9. Scatter plots of Number of intolerant fish species (nFiInt) against urban and agricultural LUs Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data). 251

Figure 6.10. Scatter plots of Number of EPT families (nEPTFa) against urban and agricultural LUs Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data)... 252

Figure 6.11. Scatter plots of stream health assessment scores against urban LU gradients. Assessment scores were produced by regional raw MMIs (top row), normalized regional raw MMIs (middle row), and normalized composite MMIs (bottom row) for fish and benthic macroinvertebrates. Blue and circles indicate Michigan samples and red and x mark symbols indicate Korean samples (large river sites excluded). 253

Figure 6.12. Scatter plots of stream health assessment scores against agricultural LU gradients. Assessment scores were produced by regional raw MMIs (top row), normalized regional raw MMIs (middle row), and normalized composite MMIs (bottom row) for fish and benthic macroinvertebrates. Blue and circles indicate Michigan samples and red and x mark symbols indicate Korean samples (large river sites excluded). 254

Abstract

As environmental degradation now reaches around the globe, ecosystem-assessment techniques and tools (EATTs) are needed in new places and at physical scales that lie outside the previous boundaries of our accumulated technical experience. To meet this need many developing and less developed countries have adapted existing EATTs from the more developed world. In this case careful evaluation is required for their suitability in a new ecological context. I refer to this issue as tool “transferability.” A related issue arises in the context of inter-regional or very large-scale assessments. Since assessments occur in specific ecoregional settings, meta-analysis of accumulating national or regional assessment datasets must be free of contextual bias inherent in statistical data gathered using different methodologies, constrained by differing geographic particularities, and reflecting the responses of locally adapted biota. This is an issue I refer to as assessment data “comparability.”

My dissertation consists of six chapters treating various issues that arise when one tries to compare ecological assessment data from two very different parts of the world: in this case Michigan and South Korea. Chapter 1 introduces general background of EATT issues and case study regions. In chapters 2-5, I analyzed transferability of hydrologic modeling, biological field sampling techniques and indicator metric development. The analysis in chapter 6, used hydrologic modeling (chapters 2 and 3) and sampling method calibrations (chapters 4 and 5) to correct regional biases in both datasets. I then used residualization techniques to correct covariate biases and directly compare the response of biological communities to urban and to agricultural land use gradients. I found (1) South Korean methods were less efficient for fish sampling but more efficient macroinvertebrate sampling; (2) methodological calibration functions were required to account for these regional differences in sampling method; (3) regional ecological normalization (residualization) and rescaling proved necessary for an unbiased comparison of LU stressor-response relationships across regions. Overall, my study suggests that EATT transferability and assessment comparability are significant but under-appreciated problems in ecological assessment and that explicit correction of regional biases are necessary for comparative analysis.

Chapter 1 : Disseratation introduction and overview

Overview of ecosystem-assessment techniques and tools

The need for ecosystem-assessment techniques and tools (EATTs) to help manage and protect existing river ecosystems and related natural resources continues to increase with growing global population, industrial development, and water demand. River ecosystems face a wide variety of anthropogenic threats including hydraulic alterations of channels for flood control, point and non-point source pollution, and radical modification of watershed hydrology due to both changing land-use and consumptive water withdrawal (Seelbach and Wiley 1996, Hughes and Hunsaker 2002, MDNR 2002, Brenden et al. 2006, Riseng et al. 2006, Allan et al. 2013). Human disturbances not only directly influence the biological diversity and community balance of natural systems (Allan et al. 1997, Lammert and Allan 1997, Wang et al. 2001, Riseng et al. 2004), but also can affect human quality of life (Bradley and Altizer 2007, Esbah 2007). Thus, around the world, decision makers and planners need to both evaluate current environmental conditions, and then develop effective strategies to protect and restore the ecological services associated with rivers (Rabeni and Sowa 1996, Higgins et al. 1999, Seelbach et al. 2002).

To quantify environmental change and degree of anthropogenic impact, many different EATTs have been developed, particularly in the US and Europe (Cairns and Pratt 1993, Karr 1995, Karr and Chu 1999, Davies 2000, Hemsley-Flint 2000, Resh et al. 2000, Verdonschot and Nijboer 2000, Wright 2000). Often developed by governmental- and nongovernmental-environmental agencies, these EATTs typically use biological indicators to simplify measurement of important ecological changes affecting both physical and biological system integrity (Merritt and Cummins 1996, Olsen et al. 1999, Fore and Yoder 2003). In fluvial ecosystem studies, periphyton, macroinvertebrate and fish assemblages are often preferred biological indicators due to their sensitivity to pollution, limited mobility, ease of collection, and relatively large quantity of taxa and individuals (Karr 1981, Hellawell 1986, Rosenberg and Resh 1993, Merritt and Cummins 1996). Macroinvertebrate and fish assemblage metrics have been used since the late 1980s throughout most of the western world as the primary tool in assessing biological changes resulting from anthropogenic impacts (Hilsenhoff 1987, Johnson et al. 1993, Resh and Jackson 1993, Fore and Yoder 2003). Similar approaches are now being applied in Asia, Africa, and much of the developing world (Moog 2007, Resh 2007, Alam et al. 2008, NIER 2009, Stubauer et al. 2010).

Global transferability of EATTs

To date, few studies have directly discussed the transferability of regionally developed EATTs to other regions of the world. Neither has there been much explicit discussion of issues of comparability in the use of these in inter-regional and global ecosystem-assessment programs (MEA 2005, Furse et al. 2006, US EPA 2006). For example, in the U. S. the Environmental Protection Agency (US EPA) both through required state-developed annual assessment reporting, and its national Environmental Monitoring and Assessment Program (EMAP), is charged with evaluating the status and trends of national ecological resources across 9 major ecoregions of the United States (US EPA 2007). EMAP was initiated because of difficulties experienced in comparing and summarizing state-level status reporting to evaluate the nation’s habitat quality and ecological integrity (Shapiro et al. 2008). Although both governmental and nongovernmental organizations have consistently have collected water quality and biological data for years (Faustini et al. 2009), differences in methodology and sampling designs made synthetic analyses (i.e., meta-analyses) difficult. In 2006 US EPA published the first national survey of US streams (The National Wadeable Streams Assessment; WSA, US EPA 2006). This report, using new data, gave a snapshot of stream condition across the nation and showed how additional standardized federal monitoring programs could provide critical information to guide resource management (US EPA 2006).

The Millennium Ecosystem Assessment (MEA 2005) recently evaluated world biodiversity using a biological indicator (total taxa richness) across 33 global sub-regions. This report evaluated the loss of biodiversity using 1970 as a reference condition for current and future scenarios in each region. Regional differences in stressor-response relationship or in trends were not compared due to difficulties in comparing regional datasets. Likewise, the European Union Water Framework Directive (EU-WFD) assessed European streams and rivers using a newly developed standardized protocol; and compared its accuracy to the (eleven) existing national rapid bioassessment programs (RBPs) used by EU member nations (Furse et al. 2006). Again direct comparisons of the national datasets were problematic. Similar but separate transnational studies in Europe have also been carried out for lakes (Lanois et al. 2011, Argilliar et al. 2013).

Each of these efforts encountered substantive difficulty in integrating existing assessment datasets from across their focal regions, and all ended up requiring new (and redundant from a public policy perspective) data collection using new standardized methodologies, or developing new standardized indicator metrics to be applied in larger scale analysis. Without careful consideration of transferability, regionally developed EATTs may lead to misinterpretations when applied in substantially different landscapes or to very different river faunas. In reality, many developing and less developed countries have not established EATTs based on their own regional conditions, but have adapted monitoring techniques and tools from developed world. Some have more or less directly adopted well-established EATTs from the developed countries without any evaluation of their transferability (MEA 2005, Furse et al. 2006, US EPA 2006, Stubauer et al. 2010). In all these cases, ensuring that results are free from geographic-and contextual biases is essential if the larger-scale analysis is to be considered valid.

I am interested in this question of the global transferability of EATTs and in how similar landscape stressors may or may not lead to similar ecological responses in very different ecological contexts (different ecoregions and/or different biological faunas). Anthropogenic stressors are known to shape many aspects of aquatic ecosystems. However, each aquatic ecosystem has a unique combination of natural geomorphology, hydrologic and hydraulic stresses, and species that may make it difficult to develop a standard set of EATTs that can apply in an unbiased way to aquatic ecosystems everywhere. If assessment results generated by EATTs can be biased by the uniqueness of regional ecosystem characteristics, then direct comparison in regional assessments will be problematic; and explicit calibration or de-biasing of regional EATT data is a necessary first step in comparative studies. Due to the global diversity of both ecological and evolutionary processes, we should at least ask how EATTs applied across eco-regional and international boundaries might be expected to vary in efficiency and bias.

There are three reasons to be cautious about concluding that EATTs are easily transferable. The first is that different ecoregions can have very different combinations of landscape features, including different variations in geology and topography, hydrological and hydraulic dynamics, and geochemical constraints (Vannote et al. 1980, Zorn et al. 2002, Wang et al. 2003, Diana 2004, US EPA 2006, Allan and Castillo 2007, Riseng et al. 2010). This global variability in geology and climate ensures that many extrinsic (exogenous) factors shaping aquatic ecosystems are relatively unique in every region. This uniqueness causes difficulty in global transferability of EATTs since 1) most have been developed or modified to be suitable for at specific regional spatial scales (e.g., state or watershed boundaries) and 2) differences in exogenous forcing leads to differences in observed correlations and their interpretation (Wright 2000, Schoolmater et al. 2013). The same is true of regional benchmark (reference) conditions used to evaluate the current status of sites within a specific region. Thus the global transferability of EATTs across major ecoregion boundaries has also been questioned based upon differing expectations for reference condition driven by different regional contexts (MEA 2005, Furse et al. 2006, US EPA 2006, Stubauer et al. 2010).

A second concern is the fact that biological species and assemblages do not just reflect the ecosystem in which they live, but also evolutionary and zoogeographic histories of the region. Given that, species-dependent biological indicators and community metrics developed for a specific aquatic ecosystem may not have the same biological meaning or sensitivity in very different biological community settings. In reality, different states, provinces, and countries have developed local indicator metrics of fish or macroinvertebrates to maximize the accuracy and efficiency of ecosystem-monitoring and -assessment programs. Regional EATT metrics usually are multi-metric indicators (MMI), Karr's IBI being a classical example (Karr 1981) developed from indicators reflecting characteristics of typical assemblages, such as functional feeding group, species richness, environmental tolerance, or sensitivity to locally common pollution regimes. Since biological community itself varies regionally, the response of biological communities to a given (even standardized) stressor need not be identical in different regions.

A third challenge, especially for the comparability of assessment results, is the need to enssure that known ecological stressors in different regions are functionally similar and comparably measured. Agricultural land use is a well-known anthropogenic stressor (Allan et al. 1997, MEA 2005, Riseng et al. 2006, US EPA 2006). However, the same proportion of agricultural land use in a catchment does not cause the same effect on aquatic ecosystems in different regions (Riseng et al. 2010). For an example, a densely populated urban region may have more intensified uses of pesticides and higher nutrient exports resulting in higher impacts on biological communities than an area with similar proportions of urban land use but with lower population densities. Land use patterns, land use intensities, and land use-related technologies are culturally mediated, and this is especially so in the cases of agriculture where resources and human market preferences cast a long shadow. Because of this, stressor-response thresholds to the gradients in land use need not necessarily be similar in different regions.

If these concerns are valid, aquatic ecosystem management and conservation should be normally regionally specified with locally appropriate EATTs. Thus, I believe we must explore how these constraints can be resolved scientifically when we ask assessment questions across and between major ecoregional boundaries.

The ecological setting of Michigan and S. Korean rivers

Throughout this dissertation I compare ecological assessment tools and data from S. Korea and Michigan regions. Here I briefly provide an overview of important regional similarities and differences. The two regions are located in globally disparate regions: S. Korea in East Asia and Michigan in North America (Figure 1.1, Table 1.1). Eventhough both regions lie in the same climate condition (temperate seasonal forest; Ricklefs 2008), they are distinguished by clearly different patterns of seasonal air temperature and precipitation, population density, and biodiversity (Figures 1.2 and 1.3, Tables 1.1-1.2, also see Chapters 2-6). Total area of Michigan (250,493 km^2) is two and half times larger than S. Korea (100,210 km^2), whereas population of S. Korea (48,661,976 people) was almost five times higher than Michigan (9,883,640) in 2010 (Table 1.1). Population density is distinctly different between the two regions: 485.6 people/km^2 in S. Korea and 67.5 people/km^2 in Michigan.

Average annual temperature in S. Korea (12.49 degree Celsius) is relatively higher than Michigan (8.22 degree Celsius) (Figure 1.2), because the latitude of S. Korea (33'06"N to 43'00"N) is relatively lower than Michigan (41'41"N to 48'18"N) (Table 1.1). Average monthly temperature showed almost similar patterns between two regions, but S. Korea is generally warmer than Michigan, for every month except December. Precipitation patterns in the two regions are distinctly different (Figure 1.3). Average annual precipitation in S. Korea (1,362mm) is much higher than Michigan (831.85mm). Furthermore, in S. Korea most precipitation is focused in a particular period of time, which includes summer monsoon season (May to September, 1,017mm), during which on average 75% of the annual precipitation falls. In contrast, Michigan has a relatively stable distribution of monthly precipitation pattern. Hydrologically the rivers of the two regions differ in terms of flow and flow yields (see Figure 6.5 in Chapter 6). At a similar size and exceedance frequency, Korean streams have on average higher flow rates and yields, indicating both higher rainfall rates (Chapter 2), higher catchment slopes (Table 6.3, Appendix 6.1 in Chapter 6) and reduced permeability reflecting the mountainous terrain and shallow soils of the interior peninsula. Due in part to these precipitation and stream flow patterns S. Korea has historically suffered from too much water in monsoon season, and too little water in other seasons; a quandary which has had a significant impact on governmental policies in regards to river management in particular, and water resources in general.

The biological communities of S. Korea and Michigan differ in substantial ways but also share influential similarities (Table 1.2). Composition of the fish fauna have little overlap at the species level (percent similarity; 4.2% of Michigan species and 3.8% of Korean species), generic (similarity; 16.0% of Michigan genera and 14.0% of Korean genera), family or order levels (similarity; 42.9% and 50.0% of Michigan and 38.7% and 64.3% of Korean, respectively). In contrast the invertebrate faunas are much more similar. At the family level faunal similarity is 76.0% of Michigan families and 77.8% of Korean families; at the order level it is 80.0% of Michigan orders and 88.9% of Korean orders.

Research goals and questions

In order to further explore and better understand these issues, I will examine transferability of EATT across Michigan streams in the USA and major river watersheds in S. Korea. Michigan and S. Korea, which have distinctly different landscapes, biology, and ecosystem research history, can provide a useful testing ground and experimental ecological comparison for exploring EATT transferability. With these biologically, morphologically, and environmentally different ecoregions, my dissertation research will investigate variations in natural and anthropogenic landscape stressors influencing biological assemblages, examine various current biological indicators and assessment techniques and tools, explore stressor-response trends to the gradients of common landscape stressors, and finally evaluate the possibility of transferability of EATT across these major ecoregions. My goal is to better understand the poorly-studied, yet ecological and practically important issue of global transferability of EATTs. Doing so could help scientifically resolve questions of global EATT transferability and help maximize the efficiency and validity of ecological assessment and management planning in both developed and developing (or less developed) countries.

Research questions that underlay my dissertation include: 1) How do differences in major landscape features and stressors affect the differing biological communities found in the rivers of the Great Lakes region and S. Korea? 2) What are the similarities and differences in useful biological indicators of the Great Lakes region and S. Korea? 3) Do biological indicators currently in use in these regions show the same stressor-response trends to gradients of land use (LU) related stressors? 4) To what extent do the differing EATTs used in these two regions complicate data comparisons and interpretation?

Dissertation content

My analyses include five research chapters based on the analysis of empirical assessment datasets either obtained from governmental agencies of Michigan and S. Korea (Chapters 2, 3, and 6) or collected from field sampling in Michigan (for fish) and S. Korea (for benthic macroinvertebrates) (Chapters 4 and 5).

In Chapter 2, I developed a series of multiple linear regression models describing the flow regime and related metrics of South Korean streams from available gauging data and catchment characteristics. I then identified key landscape variables affecting stream flow regimes in South Korea and developed a classification of the types of flow regimes that occur in South Korea. I then performed a linear modeling approach (multiple linear regression (MLR), principal component analysis (PCA), and principal component regression (PCR)) to describe, predict, and classify seasonal flow statistics from these landscape variables. Finally, the models and classification were used to estimate stream flow regimes for the un-gauged biological sampling sites used by the National Aquatic Ecological Monitoring Program (NAEMP). Stream flow regime is a influencial factor in ecological assessment because it affects stream network, channel morphology, and the distribution of biological assemblages (Poff et al. 1997, Winter 2001, Allan and Castillo 2007). However, flow regime summaries were not available for most of S. Korean bioassessment sites because of the relative paucity of stream hydrologic gauging sites. A similar lack of gauging in Michigan led to the development of modeled flow frequency statistics for un-gauged stream segments (Selbach et al. 2002) which in turn have been used in regional fisheries bioassessment studies (Riseng et al. 2006, Riseng et al. 2010, Wang et al. 2001, Wang et al. 2003). Here I developed similar models for South Korean streams to support assessment analyses in Chapter 6.

In Chapter 3, I evaluated four different analytical approaches to site-specific modeling of flow in gauged river segments. The analytical models compared included MLR, PCR, artificial neural networks (ANN), and the combination of principal components and artificial neural networks (PC-ANN). Various analytical and statistical approaches have recently been used in environmental and ecological applications with advanced technologies. The evaluation of four different models was performed to test whether MLR models in chapter 2 were a reasonable choice to predict S. Korean flow regimes or not. This chapter contributed to the discussion of the relative advantages and disadvantages of alternate methods for the estimation of site-specific stream flow regimes and regionalization of available stream gauging data.

In Chapter 4, I examined two fish sampling methodologies commonly used in rapid bioassessment programs: electrofishing and cast netting (used in Michigan and South Korea, respectively). Both Michigan and S. Korea use multimetric indicators modeled on Karr's Fish Index of Biotic Integrity (Karr 1981), but modified differently to reflect consideration of regional biology and their responses to stressors. My goal was to examine how the choice of fish sampling gear affected both sampling efficiency and metric performance. I was especially interested in what differential biases in assessment metrics could arise from the two sampling gears. Also, in this chapter I tested various fish indicator metrics in order to select a subset of individual indicator variables for the regional ecological assessments in Chapter 6.

In Chapter 5, similar to Chapter 4, I investigated potential methodological biases that might complicate comparisons of rapid bioassessment programs for benthic macroinvertebrate of Michigan and S. Korea. The Michigan Department of Environmental Quality uses a fixed-count qualitative sampling approach (e.g. 100 individuals) (MDEQ 1997). In contrast, the Korean (KNEAP) sampling uses quantitative riffle subsampling (NIER 2009). My main goal was to study how these sampling methods affected sampling performance, and the resulting potential biases in the assessment metrics. Finally, in this chapter I also conducted the comparative analysis of invertebrate datasets from two different RBPs asking how LU stressors response relationship differed between regions.

In Chapter 6, I examined issues of data comparability and integrability in the context of Michigan and S. Korean RBPs. I compared both fish and invertebrate assessment data from S. Korea and Michigan, two geographically and ecologically disparate regions, in a case study format. Specific objectives were to 1) compare Korean and Michigan ecological datasets, 2) explore the impacts of known sampling biases (Chapters 4 and 5), and regionally covarying landscape properties (Chapters 2 and 3) on their respective LU stressor-response relationships, and 3) determine the extent to which explicit corrections for methodological and statistical biases lead to altered interpretations of assessment results. Fish and benthic macorinvertebrate data from both regions were used for this study and landscape variables were summarized for each site. Regional ecological normalization (Wiley et al. 2003, Baker et al. 2005) was employed to compare LU stressor-response relationships of two regions and overall impairment rates of streams and rivers. Finally, I briefly summarized the findings of overall research and discussed the implications of this dissertation for the global transferability of ecosystem-assessment techniques and tools.

Table 1.1. Summary of general information of Michigan and S. Korea. Population densities for Michigan and S. Korea were summarized with 2010 summary data obtained from US Census Bureau 2015 and KOSIS 2015.

	S. Korea	**Michigan**
Latitude	33'06"N to 43'00"N	41'41"N to 48'18"N
Longitude	124'11"E to 131'52"E	82'7"W to 90'25"W
Total area (km^2 / mi^2)	100,210 / 38,691	250,493 / 96,716
Population	48,661,976 in 2010	9,883,640 in 2010
Pop. Density (km^2 / mi^2)	485.6 / 1,257.7	67.5 / 174.8

Table 1.2. Percent taxonomic overlap between the two regions. Genera and Species for benthic macroinvertebrates were not summarized here because MDEQ data were collected at the family level of identification. Taxa list was based on the data of National Institute for Environmental Research, S. Korea (NIER 2009) and Michigan Department of Environmental Quality, USA (MDEQ 1997).

	Classes	Orders	Families	Genera	Species
Fish					
Michigan	50.0	50.0	42.9	16.0	4.2
S. Korea	50.0	64.3	38.7	14.0	3.8
Benthic macroinvertebrates					
Michigan	100.0	80.0	76.0	-	-
S. Korea	100.0	88.9	77.8	-	-

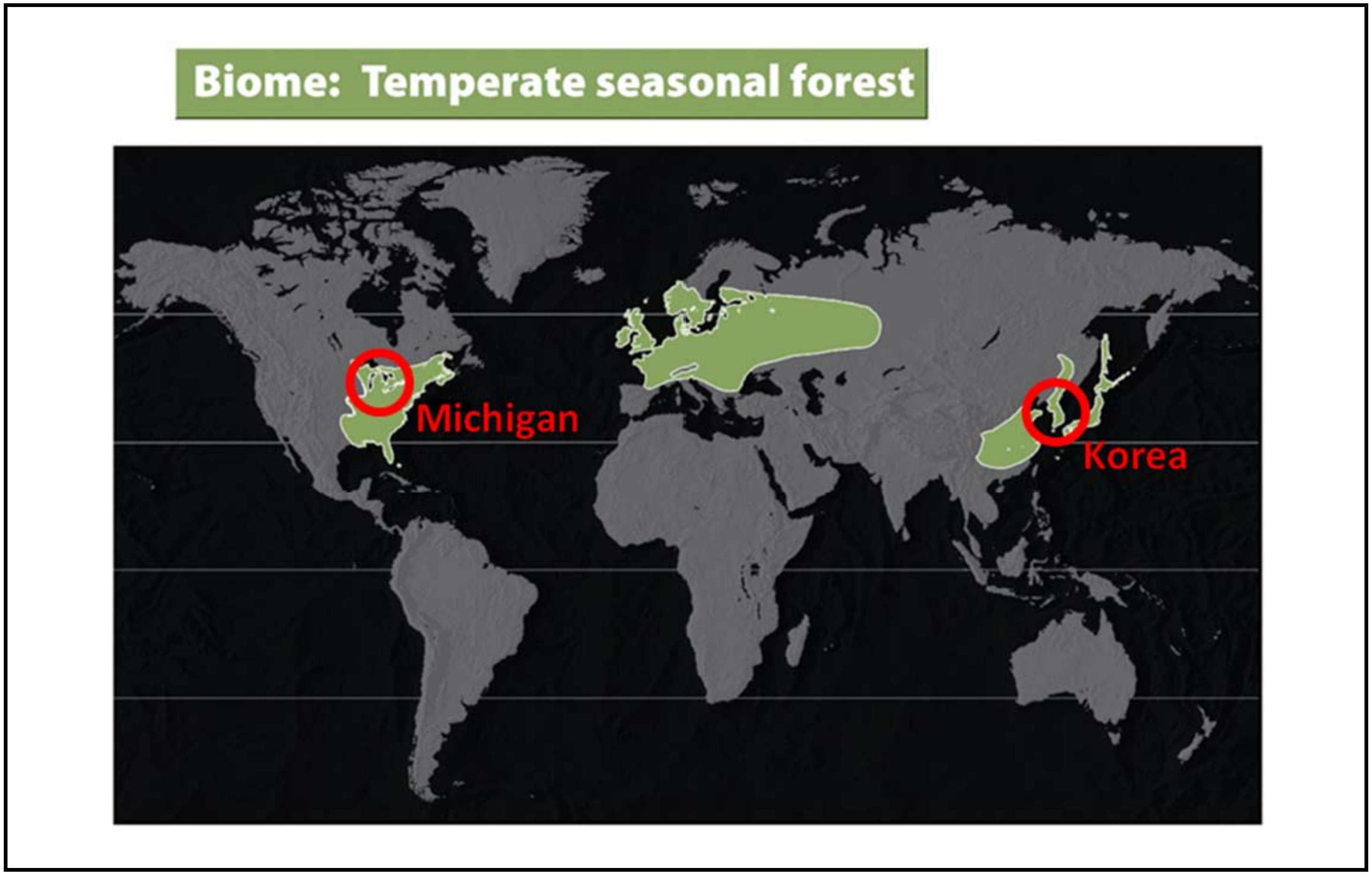


Figure 1.1. Map showing location of Michigan and Korean regions, which lies in the same ‘temperate seasonal forest’ biome. The picture is redrawn from Ricklefs 2008.

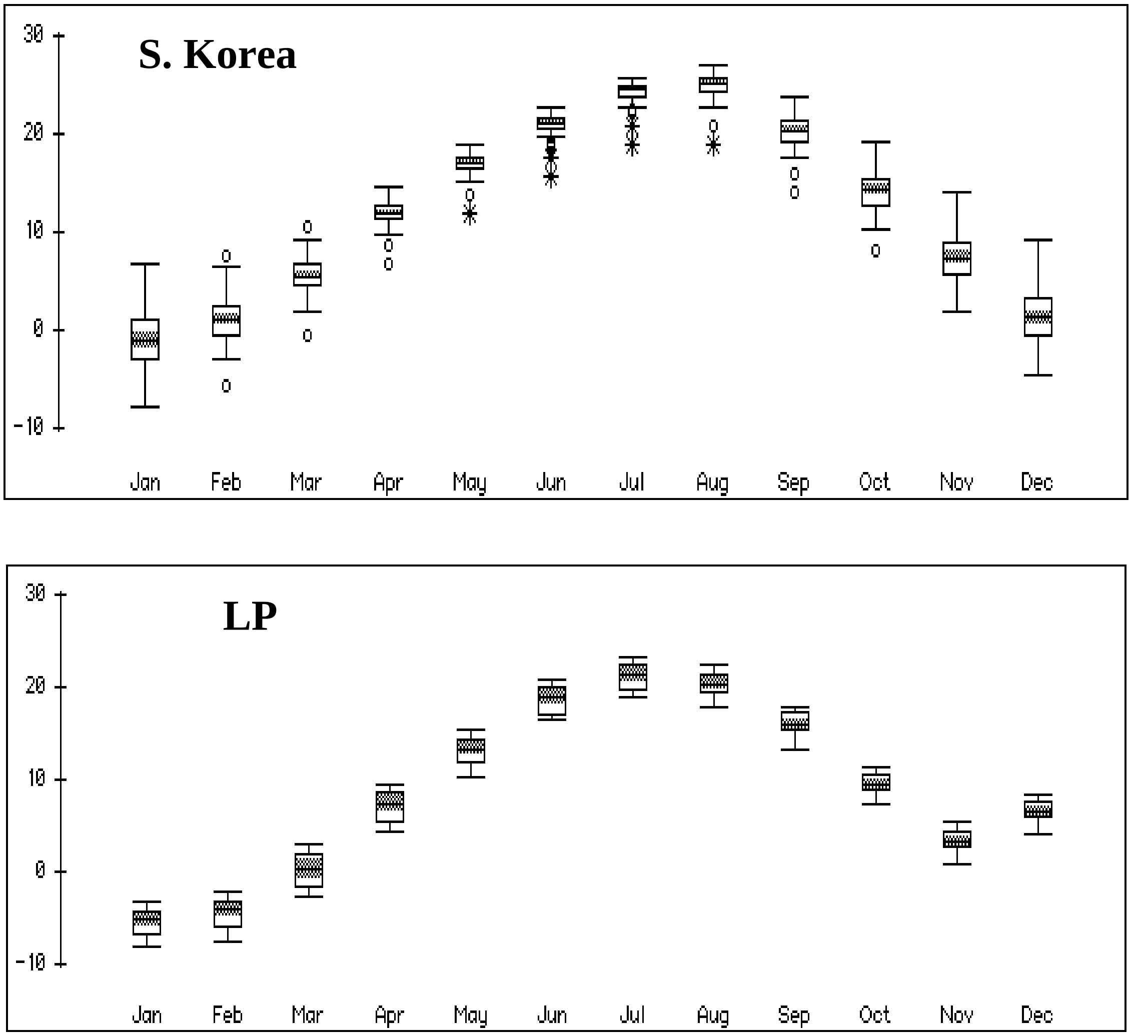


Figure 1.2. Comparison of average monthly air temperature between Lower Peninsula Michigan and S. Korea. Average monthly air temperature data for Michigan and S. Korea were summarized with data from 1981 to 2010, obtained from National Environmental Satellite, Data, and Information Service (NESDIS 2011) and Korea Meteorological Administration (KMA 2011), respectively.

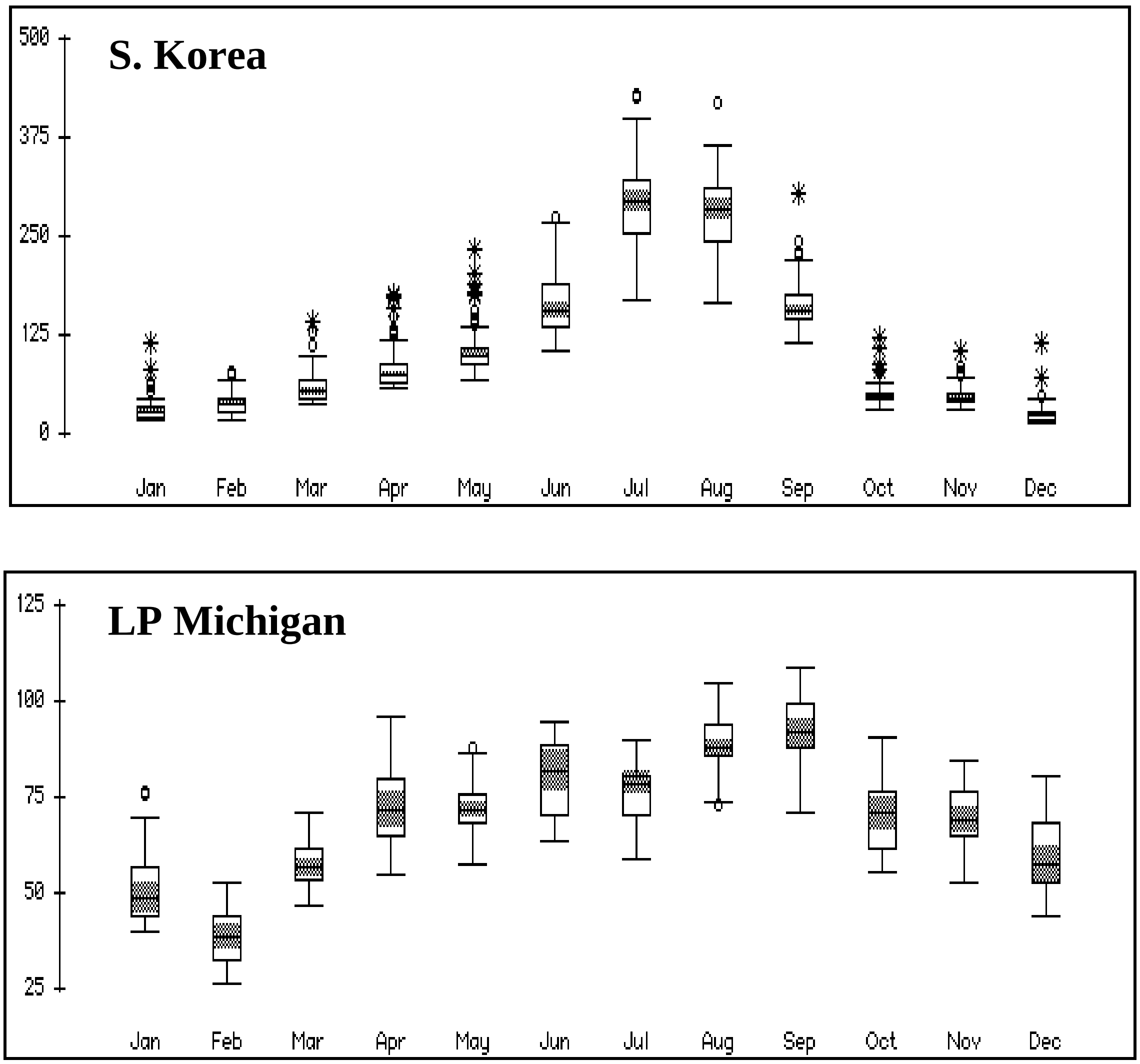


Figure 1.3. Comparison of average monthly precipitation between Lower Peninsula Michigan and S. Korea. Average monthly precipitation data for Michigan and S. Korea were summarized with data from 1981 to 2010, obtained from National Environmental Satellite, Data, and Information Service (NESDIS 2011) and Korea Meteorological Administration (KMA 2011), respectively.

Literature cited

Alam, M. S., M. M. Hoque, A. B. M. Badruzzaman, T. Huber, and B. Fliedl. 2008. Aquatic macroinvertebrates as bio-indicator: a new approach for river water quality assessment in Bangladesh. Proceedings of the Scientific Conference "Rivers in the Hindu Kush-Himalaya-Ecology and Environmental Assessment": 131-136.

Allan, J. D., D. L. Erickson, and J. Fay. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater biology 37: 149-162.

Allan, J. D. and M. M. Castillo. 2007. Stream ecology: structure and function of running waters. Second edition. Springer. Dordrecht, The Netherlands.

Allan, J. D., P. B. McIntyre, S. D. P. Smith, B. S. Halpern, G. L. Boyer, A. Buchsbaum, G. A. Burton Jr., L. M. Campbell, W. L. Chadderton, J. J. H. Ciborowski, P. J. Doran, T. Eder, D. M. Infante, L. B. Johnson, C. A. Joseph, A. L. Marino, A. Prusevich, J. Read, J. B. Rose, E. S. Rutherford, S. P. Sowa, A. D. Steinman. 2013. Joint analysis of stressors and ecosystem services to enhance restoration effectiveness. Proceedings of the National Academy of Sciences. 110 (1): 372-377.

Argillier, C., S. Caussé, M. Gevrey, S. Pédron, J. De Bortoli, S. Brucet, M. Emmrich, E. Jeppesen, T. Lauridsen, T. Mehner, M. Olin, M. Rask, P. Volta, I. J. Winfield, F. Kelly, T. Krause, A. Palm , and K. Holmgren. 2013. Development of a fish-based index to assess the eutrophication status of European lakes. Hydrobiologia. 704: 193-211.

Baker, E. A., K. E. Wehrly, P. W. Seelbach, L. Wang, M. J. Wiley, and T. Simon. 2005. A multimetric assessment of stream condition in the Northern Lakes and Forests ecoregion using spatially explicit statistical modeling and regional normalization. Transactions of the American Fisheries Society 134(3): 697-710.

Bradley, C. A., and S. Altizer. 2007. Urbanization and the ecology of wildlife diseases. Trends in Ecology and Evolution 22 (2): 95-102.

Brenden, T. O., R. D. Clark,A. R. Cooper, P. W. Seelbach, L. Wang, S. S. Aichele, E G. Bissell, and J. S. Stewart. 2006. A GIS framework for collecting, managing, and analyzing multi-scale landscape variables across large regions for river conservation and management. In R. Hughes, L. Wang, and P.W. Seelbach, (eds.), Landscape Influences on Stream Habitats and Biological Communities. American Fisheries Society Symposium 48, Bethesda, Maryland.

Cairns, J., Jr., and J. R. Pratt. 1993. A history of biological monitoring using benthic macroinvertebrates. pp. 10-27. In D. M. Resenberg and V. H. Resh (eds.). Freshwater biomonitoring and benthic macroinvertebrates. Champman and Hall, New York.

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

Diana, J. S. 2004. Biology and ecology of fishes, Second edition. Cooper Publishing Group. Traverse City, MI.

Esbah, H. 2007. Land use trends during rapid urbanization of the city of Aydin, Turkey. Environmental management 39:443-459.

Faustini, J. M., P. R. Kaufmann, and A. T. Herlihy. 2009. Downstream variation in bankfull width of wadeable streams across the conterminous United States, Geomorphology. 108 (3-4): 292-311.

Fore, L., and C. O. Yoder. 2003. The design of a biological community trend monitoring program for Michigan wadeable streams. MI/DEQ/WD-03/086. Michigan Department of Environmental Quality (MDEQ). Surface Water Quality Division. Lansing, Michigan.

Furse, M., D. Hering, O. Moog, P. Verdonschot, R. K. Johnson, K. Brabec, K. Gritzalis, A. Buffagni, P. Pinto, N. Friberg, J. Murray-Bligh, J. Kokes, R. Alber, P. Usseglio-Polatera, P. Haase, R. Sweeting, B. Bis, K. Szoszkiewicz, H. Soszka, G. Springe, F. Sporka, and I. Krno. 2006. The STAR project: context, objectives, and approaches. Hydrobiologia. 566: 3-29.

Hellawell, J. M. 1986. Biological indicators of freshwater pollution and environmental management. Elsevier, London. 546 pp.

Hemsley-Flint, B. 2000. Classification of the biological quality of rivers in England and Wales. pp 55-70. In Wright, J. F., D, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Higgins, J., M. Lammert, and M. Bryer. 1999. Including aquatic targets in ecoregional portfolios: guidance for ecoregional planning teams. Designing a geography of hope, Update 6, The Nature Conservancy, Arlington, Virginia.

Hilsenhoff, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20: 31-39.

Hughes, R.M., and C.T. Hunsaker. 2002. Effects of landscape change on aquatic biodiversity and biointegrity. pp 309-329. In Gutzwiller, K. J., editor. Applying landscape ecology in biological conservation. Springer-Verlag, New York.

Johnson, R. K., T. Wiederholm, and D. M. Rosenberg. 1993. Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. Chapman and Hall, New York.

Karr, J. R. 1981. Assessment of biotic integrity using fish communites. Fisheries. 66: 21-27.

Karr, J. R. 1995. Protecting aquatic ecosystems: clean water is not enough. In Davis, W. S. and T. P. Simson (editors), Biological assessment and criteria: Tools for water resource planning and decision making. Lewis Publishers, Boca Raton, FL: 7-13.

Karr, J. R. and E. W. Chu. 1999. Restoring life in running waters: Better biological monitoring. Island Press, Washington, DC.

Lammert, M., and J. D. Allan. 1997. Assessing biotic integrity of streams: effects of scale in measuring influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23: 257-270.

Launois, L., J. Veslot, P. Irz, and C. Argillier. 2011. Development of a fish-based index (FBI) of biotic integrity for French lakes using the hindcasting approach. Ecological Indicators. 11: 1572-1583.

Merritt, R. W., and K. W. Cummins. 1996. An introduction to the aquatic insects of North America. Kendall Hunt Publishing, Dubuque, Iowa.

Millennium Ecosystem Assessment (MEA). 2005. The millennium ecosystem assessemnt seires. Ecosystems and human well-being: Multiscale assessment. Vol. 4. Island Press. NW, Washington DC.

Michigan Department of Environmental Quality (DEQ). 1997. Qualtitative biological and habitat survey protocols for wadeable streams and rivers. Michigan Department Environmental Quality (MDEQ), Surface Water Quality Division, Great Lakes and Environmental Assessment Section, GLEAS Procedure 51 (Revised 2002). Lansing, Michigan.

Michigan Department of Natural Resources (MDNR). 2002. Huron River Plan (Livingston, Washtenaw, Oakland, Monroe and Wayne Counties). Michigan Department of Natural Resources, Fisheries Division. Lansing, Michigan.

National Institute for Environmental Research (NIER). 2009. Final reports of stream health monitoring for Korean river ecosystem. Ministry of Environment. Incheon city. South. Korea.

Moog, O. 2007. Manual on pro-rata multi-habitat-sampling of benthic invertebrates from wadeable rivers in the HKH-region. Deliverable 8, part 1 for ASSESS-HKH. European Commision, Brussels. 29 pp.

Olsen, A., R. J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Young. 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54: 1-45.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. C. Stromberg. 1997. The natural flow regime: a paradigm for river conservation and restoration. Bioscience. 47: 769-784.

Rabeni, C. F., and S. P. Sowa. 1996. Integrating biological realism into habitat restoration and conservation strategies for small streams. Canadian Journal of Fisheries and Aquatic Sciences 53 (Supplement 1): 252-259.

Resh, V. H., and J. K. Jackson. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. pp 195-233. In D. M. Rosenberg and V. H. Resh (eds.). Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Resh, V. H., D. M. Rosenberg, and T. B. Reynoldson. 2000. Selection of benthic macroinvertebrate metrics for monitoring water quality of the Fraser River, British Columbia: implications for both multimetric approaches and multivariate models. pp 195-206. In Wright, J. F., D, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Resh, V. H. 2007. Multinational, freshwater biomonitoring programs in the developing world: lessons learned from African and Southeast Asian river surveys. Environmental Management. 39: 737-748.

Ricklefs, R. E. 2008. The economy of nature. 6th Edition. W. H. Freeman and Co.

Riseng, C. M., M. J. Wiley, and R. J. Stevenson. 2004. Hydrologic disturbance and nutrient effects on benthic community structure in Midwestern U. S. streams: a covariance structure analysis. Journal of the North American Benthological Society 23: 309-326.

Riseng, C. M., M. J. Wiley, R. J. Stevenson, T. Zorn, and P. W. Seelbach. 2006. Comparison of coarse versus fine scale sampling on statistical modeling of landscape effects and assessment of fish assemblages of the Muskegon River, Michigan. pp 555-575. In R. Hughes, L. Wang, and P. W. Seelbach, editors, Landscape influences on stream habitats and biological communities. American Fisheries Society Symposium 48, Bethesda, Maryland.

Riseng, C.R., M.J. Wiley, B.E. Pijanowski, P.W. Seelbach, and R.J. Stevenson. 2010. An ecological assessment of Great Lakes tributaries in the Michigan Peninsulas. Journal of Great Lakes Research. GLR-D-09-00166R2. In Press.

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

Schoolmaster, D. R. Jr., J. B. Grace, E. W. Schweiger, B. R. Mitchell, and G. R. Guntenspergen. 2013. A causal examination of the effects of confounding factors on multimetirc indices. USGS Saff--Published Research. Paper 769.

Shapiro, M. H., S. M. Holdsworth, and S. G. Paulsen. 2008. The need to assess the condition of aquatic resources in the US. Journal of North American Benthological Society. 27 (4): 808-811.

Seelbach, P. W. and M. J. Wiley. 1996. An assessment of the potential for ecological rehabilitation and restoration in Mill Creek. Huron River Watershed Council (HRWC). Ann Arbor, Michigan.

Seelbach, P. W., M. J. Wiley, P. A. Soranno, and M. T. Bremigan. 2002. Aquatic conservation planning: using landscape maps to predict ecological reference conditions for specific waters. pp 454-478. In Gutzwiller, K. J., editor. Applying landscape ecology in biological conservation. Springer-Verlag, New York.

Stubauer. I, D. Hering, T. Korte, A. Hoffmann, K. Brabec, S. Sharma, M. Shrestha, M. A. Kahlown, M. A. Tahir, A. Kumar, M. P. Sharma, M. F. Bari, A. B. M. Badruzzaman, G. K. Chhopel, and O. Moog. 2010. The development of an assessment system to evaluate the ecological status of rivers in the Hindu Kush-Himalayan region: introduction to the special feature. Hydrobiologia. 651: 1-15.

US EPA (U. S. Environmental Protection Agency). 2006. Wadeable stream assessment: A collaborative survey of the nation's stream. Office of Water. EPA 841-B-06-002.

US EPA (U. S. Environmental Protection Agency). 2007. National water quality inventory: report to congress. EPA/841/R-07/001. Office of Water, Washington, DC.

Vannote, R. G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences. 37: 130-137.

Verdonschot, P. F. M. and R. C. Nijboer. 2000. Typology of macrofaunal assemblages applied to water and nature management: a Dutch approach. pp 241-262. In Wright, J. F., D, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Wang, L., J. Lyons, P. Kanehl, and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28: 255-266.

Wang, L., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela, and P. Stewart. 2003. Watershed, Reach, and Riparian Influences on Stream Fish Assemblages in the Northern Lakes and Forest Ecoregion, USA. Canadian Journal of Fisheries and Aquatic Sciences. Can. J. Fish. Aquat. Sci. 60:491-505.

Wiley, M. J., P. W. Seelbach, K. Wehrly, and J. S. Martin. 2003. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. pp 201-223. In T.P. Simon, editor. Biological response signatures: indicator patterns using aquatic communities. CRC Press, Boca Raton, Florida.

Winter, T. C. 2001. The concept of hydrologic landscapes. Journal of the American Water Resources Association. 37: 335-349.

Wright, J. F. 2000. An introduction to RIVPACS. pp 1-24. In Wright, J. F., D, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2002. Distributions of Stream Fishes and their relationship to stream size and hydrology in Michigan's Lower Peninsula. Transactions of the American Fisheries Society 131:70-85.

Chapter 2 : Estimation and classification of flow regimes for South Korean streams and rivers

Abstract

The information of stream flow discharge characteristics continues to be norm in watershed management and natural resource conservation, in that stream flow regime is a crucial factor influencing water quality, geomorphology, and the community structure of stream biota. The objectives of this study were to estimate Korean stream flows from landscape variables, classify stream flow gages using hydraulic characteristics, and then apply these methods to ungaged biological monitoring sites for effective ecological assessment. Here I used a linear modeling approach (multiple linear regression (MLR), principal component analysis (PCA), and principal component regression (PCR)) to describe and predict seasonal flow statistics from landscape variables. MLR models were successfully built for a range of exceedance discharges and time frames (annual, January, May, July, and October), and these models explained a high degree of the observed variation with r squares ranging from 55.5 (Q95 in January) to 89.9 (Q05 in July). In validation testing, predicted and observed exceedance discharges were all significantly correlated ($p<0.01$) and for most models no significant difference was found between predicted and observed values (Paired samples T-test; $p>0.05$). I classified Korean stream flow regimes with respect to hydraulic and hydrologic regime into four categories: flashier and higher-powered (F-HP), flashier and lower-powered (F-LP), more stable and higher-powered (S-HP), and more stable and lower-powered (S-LP). These four categories of Korean streams were related to the characteristics of environmental variables, such as catchment size,

site slope, stream order, and land use patterns. I then applied the models at 684 ungaged biological sampling sites used in the National Aquatic Ecological Monitoring Program in order to classify them with respect to basic hydrologic characteristics and similarity to the government's array of hydrologic gauging stations. Flashier-lower powered sites appeared to be relatively over-represented and more stable-higher powered sites under-represented in the bioassessment data sets. Overall, this study not only provides a straightforward and very cost-effective method to estimate stream flow discharge characteristics, but also provides a fundamental covariate data for the comparability of ecological assessments.

Introduction

Stream flow regime is a crucial factor influencing stream network and channel morphology, water quality, stream biota and instream habitats (Poff et al. 1997, Winter 2001, Allan and Castillo 2007). Stream flow variation results from a complicated set of interactions between natural setting (e.g., regional climate, geomorphic condition, and geologic material) and anthropogenic activities (e.g., land-cover alteration, channel modification, and dam construction) (Montgomery 1999, Trush et al. 2000, Diana 2004, Thorp et al. 2005, Allan and Castillo 2007). Therefore, it is widely acknowledged that river resource managers need to both document flow regimes and understand their relationships to aquatic habitats and biota (Dunne and Leopold 1978, Naiman et al. 2002, Johnson et al. 2007, Riseng et al. 2011, Seelbach et al. 2011).

In South Korea there have been growing efforts to understand and manage stream flow regimes as a part of improving public water resource policy; a pressing task given the high average population density (465 people per square kilometer in 2009; WAMIS 2013) and its unique setting in a monsoonal climate with mountainous terrain (KMA 2011). A humid continental climate with additional impacts of the East Asian monsoon leads to extremely high peak river flows and resulting sediment erosion, aquatic habitat destruction, severe flooding of densely populated regions, and degradation of water quality. In response the South Korean government has invested tremendous amounts of money to build dams and straighten major channels. To date, these efforts have been focused on larger coastal river reaches and local streams that have notable water quality issues or flow variation.

Despite a substantial investment in a national stream gauging system (602 listed stream discharge gauges) with 0.03 gauges per stream mile (NIER 2009, KMA 2011, Hwang et al. 2011, WAMIS 2013), stream flow data are still available for very few sites and basin level characterizations of hydrologic regime are lacking. Current efforts to develop sophisticated water quality and ecological monitoring programs are hampered by the lack of site-specific discharge data. For example, Li et al. (2012), in the most recent study of South Korean streams and rivers, examined the relationship between macroinvertebrates and environmental variables at multiple scales. Even though the environmental variables in this research included various geographical, land use, substratum, and physicochemical parameters, hydrological variables did not include stream discharge (volume flow rate) information. In the same manner, most nationwide ecological assessments for South Korean aquatic ecosystems (e.g., Bae et al. 2011, Cho et al. 2011a, Hwang et al. 2011, Lee et al. 2011a, Lee et al. 2011b, Yoon et al. 2011) have not dealt directly with flow variability due to the general lack of flow regime data. Since existing analyses of South Korean aquatic resources lack explicit reference to hydrologic variability (natural and/or anthropogenic), they are less than convincing in terms of understanding current environmental stresses.

Despite the acknowledged importance of stream flow regimes in monsoonal climate areas (KMA 2011, WAMIS 2013), the existence of a national gauging system, and the relative ease of modern hydrologic model development, there has been little work to date on the prediction of site-specific flow regimes for South Korean streams and rivers. Empirical MLR modeling based on assumptions of hydraulic geometry and catchment characteristics have been widely employed elsewhere to estimate stream flow and water temperature regimes for ungauged study locations (e.g., Holtschlag and Croskey 1984, Dunne and Leopold 1978, Wehrly et al. 1997, Legendre and Legendre 1998, Smakhtin 2001, Wiley et al. 2003, Allan and Hinz 2004, Hamilton et al. 2008, Seelbach 2011). For example, flow regimes for all NHD river segments across three states: Illinois, Michigan, and Wisconsin were successfully predicted as a part of a multi-agency river classification project in the Midwestern USA (Seelbach et al. 2011). In a similar way, natural flow regimes for rivers of the Great Lakes Basin were earlier assessed by Allan and Hinz (2004), and site-specific summer flows for water withdrawal permits in Michigan are currently computed from linear models of limited gauging data (Hamilton et al. 2008). Empirical modeling with landscape attributes is equally appropriate for the estimation of site-specific stream flow regimes in the South Korean peninsula, where there is currently a great need to incorporate flow information into site-based ecological assessments.

My overall objectives in this study were to: 1) develop MLR models describing the flow regime and related metrics of South Korean streams from available gauging data and catchment characteristics; 2) identify key landscape variables affecting stream flow regimes in South Korea; 3) develop a classification of the types of flow regimes that occur in South Korea; and 4) to demonstrate the use of models and classification to estimate stream flow regimes for the un-gauged biological sampling sites of the National Aquatic Ecological Monitoring Program (NAEMP).

Materials and methods

Characteristics of Korean streams and rivers

South Korean streams and rivers are largely included in five major watersheds: the Han River, Geum River, Nakdong River, Youngsan River, and Seomjin River (Table 2.1 and Figure 2.1). The total area and total stream length included in these South Korean watersheds are 109,027 km^2 and 29,809 km, respectively. The Han River Watershed is the largest watershed (41,957 km^2) followed by the Nakdong River Watershed (31,785 km^2), the Geum River Watershed (17,537 km^2), Youngsan River Watershad (12,833 km^2), and Seomjin River Watershed (4,914 km^2). However, the Nakdong River Watershed (9,637 km) has the longest total stream length among the five Korean watersheds followed by Han River Watershed (8,568 km), Geum River Watershed (6,135 km), Youngsan River Watershed (3,540 km), and Seomjin River Watershed (1,929 km). Table 2.1 summarizes other useful characteristics for each basin (Hwang et al. 2011, NIER 2009, WAMIS 2011). Discharge gaging sites of Youngsan River and Seomjin River were combined into the Youngsum River Watershed for this analysis following the watershed grouping used by the National Aquatic Ecological Monitoring Program (NAEMP), the National Institue for Environmental Research (NIER), South Korea (NIER 2009).

Data collection and summary

Discharge data for the period of record from each river gage in South Korean was obtained from the WAter Management Information System (WAMIS), NIER, Korea Ministry of Environment (WAMIS 2011). Of the 603 listed discharge gages, daily discharge data from 163 gages (Figure 2.1A) were employed in this study. I eliminated 440 sub-optimal sites, based on the following four selection criteria: First, discharge gages not operational in 2009 were eliminated because much of the biological data I am interested in were obtained in 2009. Second, in order to ensure reasonable estimates of frequency, gages with less than ten years of daily discharge data were removed with the exception of nineteen sites which were never-the-less included to balance geographic coverage, although they only had eight or nine years of discharge data. Third, gages with discharges that were heavily impacted by anthropogenic activities were removed (principally large upstream dams). Last, extreme outlier sites (11 out of 174 sites) were excluded based on boxplot and scatter plot assessments with landscape variables.

With the qualified stream discharge data from the 163 gages, annual exceedance discharges (5%, 10%, 25%, 50%, 75%, 90%, and 95%) were summarized as were similar exceedance frequencies for four seasonal time windows (Table 2.2). The analysis time windows were defined by seasonal patterns in flow variation and biological monitoring seasons: July (high flow season), January (low flow season), May (spring biological sampling season), and October (fall biological sampling season). Exceedance flows and plotted flow duration curves for each site were calculated using HEC-DSSVue 2.0.1 (U.S. ACE 2011). The smaller exceedance frequencies (i.e. the 5% and 10% exceedance flows) indicate higher flow conditions for the data series, while the larger exceedance frequencies (i.e. the 95% and 90% exceedance flows) correspond to persistent or low flow conditions (HEC-DSSVue user's manual, USACE 2011).

Candidate variables describing various landscape attributes (independent variables) to develop multiple linear regression models were summarized at the catchment scales from the digital maps of elevation, mean precipitation, mean air temperature, mean humidity, catchment slope, land cover/land use, and surficial geology (soil name and soil infiltration rate) using ArcGIS 9.1 (ESRI 2005). The digital maps of land cover/land use, surficial geology, and elevation were obtained from the WAter Management Information System (WAMIS 2011) of NIER, South Korea. The surficial geology maps included soil types and soil penetration rates. Soil penetration rates roughly reflect infiltration rates and were categorized from very excellent with high penetration rate (category 1) to very poor with very low penetration rate (category 7). Catchment slope in percentage was calculated by averaging all aspect values in percentage (ESRI 2005) and site slope was calculated by dividing elevation difference between two stream points by stream distance from the digital elevation map (Gordon et al. 2004). The digital contour maps of regional climate data (mean precipitation, mean air temperature, and mean humidity) were created with the observed data obtained from the Korea Meteorological Administration (KMA 2011), Rep. of Korea (Figure 2.1B). The regional climate data includes mean annual summaries collected from 1981 to 2010 at 63 operational weather observation stations. The final list of candidate landscape attributes included catchment size, latitude, longitude, altitude, catchment slope, channel slope, mean precipitation, mean air temperature, mean humidity, number of dams, proportions of land-use type, proportions of soil type, and proportions of soil infiltration rate (Table 2.3).

Multiple linear regression model development

MLR models were constructed in order to predict a series of stream exceedance discharge frequencies (5%, 10%, 25%, 50%, 75%, 90%, and 95%). Initial selection of predictor variables in the model was based on previous research that had identified important environmental factors influencing stream exceedance discharges in the Midwestern USA (Wiley et al. 1997, Allan and Hinz 2004, Hamilton et al. 2008, Seelbach et al. 2011). If necessary, the dependent and independent variables in MLR models were transformed to natural log form after adding the integer 1 or 0.01 to the variable in order to maximize linearity within the modeled relationships and to meet assumptions of normality for all variables (Wiley et al. 2003, Riseng et al. 2006). To be specific, 1 was added to catchment area, precipitation, and site elevation and 0.01 was added to catchment slope, site slope, land cover/land use, and surficial geology. However, number of dams was not transformed into natural log form in the model because the variable showed better linearity and normality without transformation.

The addition of independent variables in the MLR models to predict stream exceedance discharges was carried out using a manual, stepwise regression approach (Hocking 1976, Draper and Smith 1981) using Datadesk 6.0 (Velleman 1997). Independent variables were inserted in the model in the following order: 1) catchment area, 2) mean precipitation, air temperature, and humidity, 3) catchment slope or site slope, 4) site elevation, 5) land cover/land use, 6) surficial geology variables (soil type or penetration rate), and 7) number of dams. Independent variables were retained in the MLR models that would maximize R^2, be significant at $p < 0.05$, and have a T-ratio greater than 2 in the model (Wehrly et al. 1997, Wiley et al. 1997, Wiley et al. 2003, Riseng et al. 2010, Seelbach et al. 2011). In several models, exceptions occurred in which I included variables with T-statistics $p < 0.10$, because they appeared critical to maintaining high R^2 for the prediction. If two independent variables in the model were significantly correlated, only the variable that best improved model fit was retained.

The stream discharge gauging dataset was randomly partitioned into model-building and model-testing groups. A total of 30 gages, approximately 22.6% of the total gages, were set aside as a model-testing group and were used to evaluate and validate models. The remaining 133 sites (approximately 77.4%) were set into a model-building group, which was used for building MLR models for each exceedance discharge.

Performance evaluation of four different predictive models

In order to evaluate the performance of each predictive model, the Mean Absolute Error (MAE) and Nash-Sutcliffe model efficiency (NSE) coefficient were computed with the observed and predicted flow discharges of each percent exceedance freequency. The MA) is a statistical approach used to measure how close predicted values are to the observed values and can be defined as follows,

$$\text{MAE}=\left[n^{-1}\sum_{i=1}^{n}\left| x_{obs}-x_{pre} \right| \right] \tag{1}$$

where n indicates the number of observations of stream discharge for each percent exceedance (Hyndman and Koehler 2006). Here, x_{obs} and x_{pre} indicate the observed and predicted stream discharges, respectively. The Nash-Sutcliffe model efficiency coefficient (NSE) was also computed to evaluate the predictive power of each exceedance model using the predicted and observed stream discharges (Nash and Sutcliffe 1970). If NSE value is greater than 0.5, the model shows acceptable accuracy. If NSE value is greater than 0.7, it the model is in a good agreement with observation (Moriasi et al. 2007). The NSE can be defined as follows,

$$\text{NSE}=1-\frac{\sum_{t=1}^{T}(x_{obs}^{t}-x_{pre}^{t})^{2}}{\sum_{t=1}^{T}(x_{obs}^{t}-\overline{x}_{obs}^{t})^{2}} \tag{2}$$

where x_{obs} is the observed stream discharges, x_{pre} is the predicted stream discharges, and $\overline{x}_{obs}^{t}$ is the averaged value of the observed stream discharges.

Korean stream flow-type classification methods

A simple classification system for Korean stream flow regimes was developed from the entire set of gauged stations (n=163) using Principal Component Analysis (PCA). PCA has often been used as an indirect ordination technique to describe the main dimensions of variation in multivariate data sets (Maceda-Veiga and Sostoa 2011). PCA produces synthetic functions which are linear combinations of the original data. Kaiser's rule was used to evaluate PCA axes (Legendre and Legendre 1998, Meador and Carlisle 2007, Maceda-Veiga and Sostoa 2011). I used PCA to combine three aspects of flow variability into two synthetic ordination axes: specific stream power (Q10×site slope/wetted width; Bagnold 1966), baseflow yield (Q90/catchment area; Zorn et al. 2004), and flow flashiness (Q10/Q90; Seelbach et al. 2011). These three variables were chosen to provide a balanced representation of biologically important differences between stations in high flow hydraulic energy dissipation, low flow habitat quality, and annual flow variability, respectively. Size related variables (e.g. catchment area, link number) were purposefully excluded to ensure that variance partitioning would be constrained to differences in hydrographic pattern and not absolute flow magnitude. Prior to the ordination analysis, the data matrix was log-transformed to improve the assumption of linearity and standardized (Legendre and Legendre 1998, Wiley et al. 2003, Riseng et al. 2010, Seelbach et al. 2011, Maceda-Veiga and Sostoa 2011). The classification of Korean stream flow gauging sites was based on quartile of occurrence in ordination space, which can be computed from values of the original three variables and the axis loadings. This classification was then applied to un-gaged biological sampling sites (n=684) to evaluate the representation of streams types in the biological survey data set.

Statistical analysis

Paired samples t-test, oneway-ANOVA test, oneway-ANOVA Tukey test, and Pearson correlation were used (SPSS, Inc. 2003) to compare exceedance discharges over the annual and seasonal time windows and to compare observed and predicted exceedance discharges. Statistical summaries (mean, median, standard deviation, minimum, and maximum), box plots, and scatter plots, MLR analyses and PCA analysis were conducted using Datadesk 6.0 (Velleman 1997) and SPSS 12.0 (SPSS, Inc. 2003).

Results

Observed flow exceedance frequencies and duration curves varied geographically and seasonally (Table 2.2 and Figure 2.2) and across four major watersheds examined (Table 2.4 and Figure 2.3). Overall, measured stream flows in South Korean rivers ranged from a maximum of 7,731 cms (Q05 from Han River) to a minimum of zero (Q90 and Q95 from Han River). July had the highest mean stream discharge at each exceedance frequency, and January had the lowest mean discharges (Table 2.2 and Figure 2.2). Seasonal differences in discharge were statistically significant for Q05, Q10, Q25, and Q50 ($p<0.01$, One way-ANOVA test), whereas the low flows (Q75, Q90, and Q95) did not show statistically significant seasonal differences ($p>0.05$, Oneway-ANOVA test). Comparing the four major watersheds, mean stream discharges of the Han River Watershed were the highest at most exceedance frequencies (Q25, Q50, Q75, Q90, and Q95), although the Nakdong River Watershed showed the highest Q05 and Q10 discharges (Table 2.4 and Figure 2.3).

Individual gauging sites varied widely in drainage area, mean annual precipitation, land use, and catchment slope (Table 2.3). Mean drainage area for all sites was 3,080.42 km^2 and ranged from 50.23 km^2 to 23, 316.70 km^2. Mean average annual precipitation was 1,320 mm and varied from 1,073 mm to 1,588 mm. Urban land use was important in MLR modeling and ranges were 0 % ~ 48.8 %. Average catchment slopes varied from 0.0920 to 0.5424 and the mean was 0.3066. Number of dams above a site varied from 0 to 11 with a mean of 1.48.

MLR models

Multiple linear regression models successfully developed for all (Q05, Q10, Q25, Q50, Q75, Q90, and Q95) exceedance flows explained a high degree of the observed flow variation

(Table 2.5). R-squares of models ranged from 55.5 (Q95 in January) to 89.9 (Q05 in July) and averaged 76.8. July models had the highest average r-square (87.3) followed by annual (79.6), October (75.4), May (75.3), and January (66.4) models. Q75 and Q50 models generally had better fits (average r-squares of 82.0 and 80.9, respectively). Q95 and Q05 models generally had the poorest fits (average r-squares of 73.5 and 74.8, respectively).

In the regression models, drainage area, mean annual precipitation, catchment slope, urban land use, surficial geology, and number of dams were the key independent variables predicting discharge in South Korean streams and rivers (Table 2.5). As expected, drainage area was consistently an essential and powerful predictor of flow in all MLR models. Mean annual precipitation and urban land use were included in most of the models although they were not important in Q90 and Q95 models for January. Urban land use was likewise important in many of the models but did not contribute to models for the Annual Q05 and Q10, for Q90 and Q95 in January, and for Q05 and Q10 in July. Number of dams and surficial geology (either in the form of soil type or soil penetration rate) were also often important variables in the MLR models. Catchment slope was significant for some specific exceedance flows and time window (Q50 and Q75 in July). No significant effects of mean air temperature (ºC), mean annual humidity (%), latitude, longitude, and channel slope were detected when building MLR models (Multiple linear regression, $p>0.05$).

Evaluation of model performance

Mean Absolute Error (MAE) values generally indicated good prediction of exceedance flows in both model generation and validation steps (Table 2.6). MAE values of models ranged from 29.2 (Q05 in July) to 85.6 (Q05 in January) with average of 52.3 for training step and ranged from 33.3 (Q05 in July) to 95.2 (Q05 in January) with average of 52.6 for validation step.

In model generation step, July models had the lowest average MAE values (37.2) followed by annual (48.2), October (55.8), May (56.0), and January (64.3). Validation step also showed a similar pattern, although the average MAE (54.0) of May models was slightly lower than the average (54.8) of October models. Of all exceedance frequencies, Q75 had the lowest average MAE values (46.0 and 44.0) in both model generation and validation steps, respectively. However, the Q10 models had the highest average MAE value (57.7) in model generation step and Q25 models had the highest average (62.2) in the validation step.

Nash-Sutcliffe model efficiency (NSE) coefficients also indicated good predictive power for all exceedance models (both model generation and validation steps) (Table 2.6). NSE coefficients ranged from 0.56 (Q95 in January) to 0.90 (Q05 and Q10 in July) with average of 0.77 for training step; and ranged from 0.47 (Q10 in January) to 0.90 (Q05 in July) with average of 0.74 for validation step. For both the annual and the four seasonal time windows, July had the highest average NSE coefficients (0.87 and 0.83) followed by annual (0.80 and 0.78), October (0.75 and 0.72), May (0.75 and 0.70), and January (0.66 and 0.68) in generation and validation steps, respectively. Of all exceedance discharge frequencies, Q75 had the highest average NSE values (0.82 and 0.82). However, Q25 and Q95 had the lowest average NSE values (0.74 and 0.74) in model generation step and Q25 had the lowest average (0.64) in validation step.

Predicted and observed exceedance discharges from all sites combined were all significantly correlated each other ($p<0.01$; Table 2.7) indicating good agreement between modeled and observed values; and test and validation group correlation values for each time window did not show any significant differences (Paired samples T-test, $p>0.05$). Correlations ranged from 0.45 to 0.96 with 33.7% of the correlations above 0.90 and 78.9% above 0.80. For all S. Korean sites combined and in the four major watersheds, mean correlation for the wettest season (July) were much higher than those in other time windows, whereas correlations for the lowest flow season (January) time period were generally lower (Table 2.7). Correlation values for the Geum, Han, and Nakdong River Watersheds ranged from 0.71 (Q95 in October, the Geum River Watershed) to 0.96 (Q05 in July, the Nakdong River Watershed). However, the Youngsum River Watershed had lower correlations with values ranging from 0.45 (Q05 in January) to 0.86 (Q75 in July). In most models, higher exceedance flows had a tendency to be underestimated and lower exceedance flows to be overestimated (Figures 2.4 throughout 2.8).

In most cases, no significant difference was found between predicted and observed exceedance discharge values (Paired samples T-test; $p>0.05$). However, a significant difference was observed in Q05 and Q10 for annual and July models of the Han River Watershed when all sites were combined.

Classification of Korean stream flow types

The PCA produced two significant axes (eigenvalues >1), which explained 93.7% of variation in the input data matrix (Table 2.8). PC1 accounted for 59.5% of the variation with eigenvalue of 1.786 and was heavily influenced by baseflow yield [cms/km2] and flow flashiness (ratio). PC2 explained 34.2% of variation with eigenvalue of 1.025 and with heavier loading by specific stream power [kW m^{-2}]. Linear models for each axis were:

$$PC1 = -0.091146417\times\log(\text{specific stream power}) - 1.0607645\times\log(\text{baseflow yield}) + 0.96745328\times\log(\text{flow flashiness}) \qquad (3)$$

$$PC2 = -1.6094319\times\log(\text{specific stream power}) - 0.15671417\times\log(\text{baseflow yield}) - 0.33408582\times\log(\text{flow flashiness}) \qquad (4)$$

A scatter plot of site PC scores showed that variations in Korean stream flow regimes are widely distributed across flow stability (PC1) and stream power (PC2) gradients (Table 2.8, Figure 2.9). Based on PC1 scores for each gauging site, flow types were classified as flashier streams (PC1 site score > median of PC1 scores) or more stable streams (PC1 site score < median of PC1 scores). Using PC2 scores each site was also classified as either a higher-powered stream (PC 2 site score > median of PC2) or lower-powered stream (PC 2 site score < median of PC2). Thus, final Korean stream flows were categorized into four different flow types (Figure 2.10) based on the two-dimensional PCA ordination; flashier and higher-powered (F-HP) streams, flashier and lower-powered (F-LP) streams, more stable and higher-powered (S-HP) streams, and more stable and lower-powered (S-LP) streams.

Box plots of water chemistry data by stream classification types showed significant and consistent relationships in most of water chemistry parameters for stream discharge gages (Figure 2.11). More stable stream sites generally had higher values in catchment size, stream order, baseflow yield, and urban land use than flashier stream sites, whereas flashier stream sites had higher numbers in Q10 and Q90 ratio and proportion of forest land use. Higher-powered stream sites had higher values in site slope and stream power than those of lower-powered stream sites.

This stream classification showed that overall gaging sites were almost equally distributed in each axis, although each major watershed had slightly different patterns (Table 2.9 and 2.10, Figures 2.12 and 2.13). The Geum and Yeougsum River Watersheds had a flashier and lower-powered hydrologic regime, while the Han River and Nakdong River basins had regimes with more stable and higher specific stream power dissipation rates.

Stream flow-type classification of nation-wide biological sampling sites in S. Korea

Stream flow types of nation-wide biological sampling sites (n= 684) in S. Korea were classified using Equations 3 and 4 above, and the relative frequency of types was compared (Tables 2.9 and 2.10, Figures 2.12 and 2.13) to gauging sites and among the four major river basins. There were substantive differences between the gauged and biological sampling sites, and between the distributions of types across the 4 major basins. The classification results showed that most of nation-wide biological sampling sites (n= 684) were located in flashier streams (n= 443, 64.77%) as opposed to more stable streams (n= 241, 35.23%). Also, lower powered streams (n= 400, 58.48%) were much more frequently selected than higher powered streams (n= 284, 41.52%).

Discussion

South Korea is located in the East Asian monsoon region, and has high seasonal variation in precipitation (KMA 2011, WAMIS 2013). Hence, flow rates in South Korean streams experience dramatic seasonal changes and this has affected the stream channel morphology, types of instream substrates, and habitat conditions for stream biota (Bae et al. 2011, Cho et al. 2011, Lee et al. 2011b, Li et al. 2012). Since ecological responses of stream biota are very sensitive to seasonal changes of stream hydrology (Dudgeon. 2000, Riseng et al. 2004, Stevenson et al. 2006, Allan and Castillo 2007, Baker and Wiley 2009), seasonal stream flow estimates produced by my MLR models for the 684 NIER biological assessment sites will be very helpful in understanding geographic and seasonal variations in stream biology and health (see Chapter 5).

The explanatory power of all stream flow models was generally quite good. Models of the high flow season (July) outperformed other time periods, suggesting that catchment-scale landscape variables worked best to explain the runoff variability under saturated conditions as opposed to flows strongly influenced by groundwater, impoundment or other routing influenced by storage. In contrast, exceedance models for the low flow season (January) performed relatively poorly, although the R^2 of the low flow season still ranged from 55.5 to 78.8. Also, across all time frames, MLR models of lower discharge (e.g. Q95) also showed relatively low levels of fit with average R^2 of 73.5. These patterns are very similar to fit variations reported from the Great Lakes region (Wiley et al. 1997, Smakhtin 2001, Kilgour and Stanfield 2006, Hamilton et al. 2008, Seelbach et al. 2011) where baseflow yields were also more difficult to predict. Baseflow variation in dry periods depends strongly on subsurface routing and storage, and these processes are likely more influenced by details of local physiography than catchment scale average conditions (Baker et al. 2003).

Factors controlling stream flows in S. Korea

The regression coefficients used in this model often varied progressively in sign and weight across exceedance discharges and time windows, reflecting previously reported relationships between catchment character and stream flows (Holtschlag and Croskey 1984, Hamilton et al. 2008, Seelbach et al. 2011). In general, drainage area, mean annual precipitation, catchment slope, and site elevation had positive effects in most of the models.

A relatively strong impact of surficial geology on flow regime has been reported for Lower Michigan Rivers (Seelbach et al. 2011). In this study, surficial geology, summarized by soil penetration rate, showed very interesting relationships to discharge in most exceedance discharge models. The highest soil penetration rate class was strongly associated with high flows (Q05 and Q10) in annual and July models. However, soils with lowest penetration rate showed strong relationships on low flows (e.g., Q75, Q90, and Q95). This means that lower soil penetration rate significantly influences on stream flows in relatively dry condition, whereas higher soil penetration rate is considerable for stream flows in high flow events. However, recent studies of Korean aquatic ecosystems using environmental variables (e.g. Li et al. 2012) have not considered either soil penetration rates or soil types.

The importance of land use/land cover attributes has been well described in various stream models and ecological assessment studies (e.g., Leopold 1968, Simmons and Reynolds 1982, Hall et al. 2001, Anonymous 2003, Allan 2004, Baker et al. 2005, Brenden et al. 2006, Johnson et al. 2007, Seelbach et al. 2011, Li et al. 2012). I found several significant relationships between stream discharge regime and land use. Most of my MLR models had strong positive influence of urban land use on stream flow rate (Table 2.5). Interestingly, agricultural land use had statistically significant influences in many discharge models; however, it was dropped from my analyses because urban land use exhibited better fits and statistical power in most models and either one was not significant when both them were added in the models. In general regression coefficient signs and relative significance agree with the findings of many of previous stream studies (e.g., Werhly et al. 1997, Stauffer et al. 2000, Hall et al. 2001, Morley and Karr 2001, Riseng et al. 2006, Wang et al. 2006, Seelbach et al. 2011, Li et al. 2012).

Another interesting landscape factor useful in predicting stream discharge was the number of dams. It had positive relationships with stream discharge for lower flows, indicating that this variable is more influential on stream base flow than on higher seasonal flows.

Classification of Korean stream flow types

The climate of South Korea includes a humid continental climate and a humid subtropical climate and is also affected by the East Asian monsoon, which means that heavy precipitation is observed in a short rainy summer season and extremely cold temperature with minimum precipitation is observed in winter. These patterns were well described in the comparison of exceedance discharges in four different seasonal time windows. Also, South Korea is a mountainous peninsula, located in the middle latitudes of the Northern Hemisphere and on the east coast of the Eurasian Continent. In particular, the eastern region of South Korea has high mountain ranges and narrow coastal plains. Therefore, most upstream catchments are relatively small with higher catchment slope, resulting in flashier but lower-powered stream conditions. In contrast the western region consists of broad coastal plains, larger river basins, and rolling hills. This geomorphologic condition creates higher flows with more seasonally stable stream flow regimes. Thus, four different classes of Korean stream flow-types reflect unique regional combinations of climate and geomorphology.

My classification results showed that most biological sampling sites are located in flashier and lower-powered streams (Tables 2.9 and 2.10, Figures 2.12 and 2.13). The Ministry of Environment launched a nation-wide watershed monitoring project from 2008 and 720 biological monitoring sites were selected in 2009 (NIER 2009). The selection of the monitoring sites was mainly based on stream order, land use, proximity to a gauging location, site accessibility, importance to human life, and spatial distribution. However, flow related characteristics were not high priority in the selection of the biological sampling sites, even though some of landscape variables indirectly reflect stream flow characteristics.

Assumptions and limitations

MLR models have been applied to the estimation of various environmental factors and ecological reference conditions from spatial-scale landscape attributes for several decades (e.g., Holtschlag and Croskey 1984, Wehrly et al. 1997, Wiley et al. 2003, Moore et al. 2004, Ries et al. 2004, Wehrly et al. 2006, Hamilton et al. 2008, Seelbach et al. 2011, Cho et al. 2011b). My MLR models for South Korea provide useful estimates of annual and seasonal exceedance flows based on site-specific landscape criteria. While this information is useful as a description of expected flow regime, it does not describe year to year variation or short-term variability in Korean stream flows. MLR models have certain limitations when used to predict stochastically influenced stream flow events, because the MLR modeling approach is relatively insensitive to extreme variations (Seelbach et al. 2011, Cho et al. 2011b). Direct comparisons of MLR to other modeling approaches (e.g. see my chapter 3 for comparison to artificial neural network and principle component regression, Steen et al. 2006, Cho et al. 2011b) suggest that despite this limitation, MLR generally performs similarly to other estimation approaches and seems adequate for my purposes in Chapter 5.

The MLR models developed here might miss or ignore some important factors that directly or indirectly affect Korean stream flows which could increase their accuracy. For example, I did not use any riparian summaries of landscape variables in my models, which could be more influential on local stream flow recharges and withdrawals than catchment-scale variables (Wang et al. 2003, Wehrly et al. 2006, Zorn et al 2008, Riseng et al. 2010). In addition, anthropogenic impacts might be more critical in low flow seasons or low exceedance percentages with poorer predictive powers, because South Korea is a country with high population density and high off-channel water demands. These types of impacts might be better explained with other localized water use metrics, such as wastewater treatment discharges, residential demands for drinking water, and groundwater pumping (Winter et al. 1998, Dudgeon 2000, Bunn and Arthington 2002, Stevenson et al. 2006, Riseng et al.2010).

Although MLR models used for this study were successfully employed in the estimation of Korean stream flows, some estimation error was unavoidable in MLR models. My results showed that very high and very low flows suffered from higher prediction errors (Figures 2.4 throughout 2.8). These problems can be related to the location and number of stream flow gages used for these MLR models. In general, Korean stream flow gages are located on bigger streams or developed areas in major Korean watersheds. Thus, the number of stream gages is relatively low for smaller catchments with very low flows and for extremely large catchments with very high flows. In these cases, landscape variables could well have been relatively homogeneous and not very useful, a problem (Seelbach et al. 2011) discussed in the stream flow estimation for all rivers across three Midwest states.

Spatial scales of environmental data analysis have often been problematic in the prediction of environmental, biological, and ecological variables (Wiley et al. 2003, Riseng et al. 2006, Riseng et al. 2011, Seelbach et al. 2011). Korean stream flows can also be highly influenced by processes operating at different scales; for example different patterns of land use, geology, human culture, and various micro-scale conditions of catchments. Obvious differences in the comparison of predicted and observed stream flows for major watersheds in Korean rivers and streams were observed with the Pearson correlation tests (Table 2.7). There also appeared to be some basin specific issues. For example the combined Youngsan and Sumjin River watersheds (referred to "Youngsum R." in this study) are located in the southern part of South Korea and have a river network complicated by very complex land development patterns and multi-purpose dams. This unique environmental setting appeared to reduce the accuracy of estimation and might need to be analyzed in more detail than other major Korean watersheds.

Management implications

In this study, I emphasized the relative ease and efficiency of MLR modeling approach to estimate metrics needed to describe the hydrologic regime of South Korean streams and rivers. Since the preservation and management of water quality and aquatic ecosystems are rapidly raising concerns in many countries (Sowa et al. 2007, Zorn et al. 2008, Lee et al. 2011b, Li et al. 2012), watershed managers and planners cannot ignore the rising costs of field-based measurement, monitoring, and assessment (Wiley et al. 2003, Park 2007, Seelbach et al. 2011). Although flow regime is one of the most critical environmental factors in watershed study and management, the field monitoring of stream hydrology is neither simple nor inexpensive. In every country stream flow gages are necessarily limited in number and cannot cover all streams necessary for the wide ranges of stream type addressed in management and preservation plans.

Thus, the need to model flow regimes from limited representative data sets will continue to be the norm, and continued explorations of modeling efficiency and effectiveness will be important for effective water resource management.

My flow regime classification for Korean streams and rivers provides a useful lens for examining the selection criterion used to choose the national biological monitoring sites. Nation-wide stream health monitoring and assessment should necessarily sample all types of stream habitats in proportion to their occurrence. I found discrepancies between the relative representation of types in the national hydrology and biological assessment data sets, which raise a question about potential sampling biases. Given that there is general agreement that stream flow characteristics are essential to aquatic conservation planning and science (Petts et al. 1999, Power et al. 1999, Anonymous 2002, Ries et al. 2004, Riseng et al. 2004, Higgins et al. 2005, Piggott and Neff 2005, Seelbach et al. 2006, Sowa et al. 2007), biased selection may lead to inaccurate conclusions and redundant costs and labor. My flow estimation and regime classification approach could be used to develop a more objective method for allocating monitoring and assessment sites in S. Korea. The models could be used to estimate the actual representation of hydrologic types across the country, and assessment sampling allocated accordingly. To achieve a proportional allocation (or even objective comparison of current sampling distributions to the actual occurrence of stream types) will, however, require a more developed segment-based GIS representation of the Korean drainage system; something similar in structure to the U.S. NHD system (Anonymous. 2002).

In conclusion, this study provides a straightforward and very cost-effective method to estimate stream flow discharge characteristics and classify flow regimes in South Korea using existing summaries of common landscape variables. It is my hope that these estimates can contribute to more accurate and trustworthy environmental impact assessment, fisheries management decision making, and water supply planning.

Table 2.1. Summary of watershed characteristics of the five major river watersheds and site numbers used for the estimation and classification for stream flow regimes of stream flow gages and un-gauged biological sampling sites used by the National Aquatic Ecological Monitoring Program (NAEMP) throughout South Korea.

Watershed	Area (km^2)	Length of main stream (km)	Total stream length (km)	Number of tributaries	Human population	Number of water gauges	Number of biological sites
Geum River	17,537.0	393.1	6,134.9	876	6,205,038	38	130
Han River	41,957.0	560.0	8,567.7	912	27,046,430	43	284
Nakdong River	31,785.0	470.0	9,637.6	1,185	13,211,817	47	130
Youngsan River	12,833.4	117.7	3,540.4	576	3,004,860	23	76
Seomjin River	4,914.3	211.9	1,928.8	283	1,192,945	12	64
Total	109,026.7	1,752.7	29,809.4	3,832	50,661,090	163	684

Table 2.2. Summary statistics of stream discharges (cms) at each percent exceedance frequency in Korean streams and rivers used for normalizing linear regression models as dependent variables. Stream discharge is represented by "Q" and defined as 05-95% exceedance discharges (cms). "n" indicates the total number of gages for the summary of stream discharges and "SD" indicates standard deviation.

Dependent variable	Percent exceedance (Q)	n	Mean	Median	SD	Min	Max
Annual	05	163	331.31	84.55	676.77	4.35	5,253.50
	10	163	207.48	44.90	461.87	1.69	3,820.90
	25	163	95.27	17.00	234.19	0.64	2,219.00
	50	163	55.53	6.55	176.71	0.18	1,857.80
	75	163	33.28	3.11	108.43	0.06	1,102.90
	90	163	19.22	1.58	52.26	0.03	323.90
	95	163	14.53	1.00	41.29	0.01	285.70
January	05	163	163.35	24.63	558.21	0.39	5,253.50
	10	163	111.52	17.31	377.01	0.33	3,820.90
	25	163	63.04	8.00	196.38	0.15	2,096.20
	50	163	43.74	4.00	161.50	0.06	1,782.30
	75	163	24.93	2.13	78.79	0.03	723.00
	90	163	16.01	1.09	49.31	0.00	356.80
	95	163	11.97	0.81	39.49	0.00	341.80
July	05	163	814.89	225.16	1,451.96	10.34	7,730.90
	10	163	529.91	131.41	979.25	6.40	5,253.50
	25	163	253.53	56.92	482.34	3.72	2,612.40
	50	163	113.34	20.00	250.73	0.99	1,906.20
	75	163	58.84	8.00	153.40	0.42	1,317.20
	90	163	32.04	3.70	77.87	0.09	528.30
	95	163	23.71	2.89	60.49	0.04	415.10
May	05	163	215.17	51.52	442.23	2.68	3,823.80
	10	163	165.87	35.27	404.65	1.36	3,820.90
	25	163	93.48	15.93	225.53	0.67	2,151.90
	50	163	59.79	7.09	180.30	0.18	1,875.40
	75	163	35.96	3.58	108.05	0.06	1,047.90
	90	163	19.83	1.82	52.21	0.02	300.70
	95	163	15.37	1.23	44.01	0.01	272.60
October	05	163	173.67	38.00	433.21	0.50	3,820.90
	10	163	135.11	27.04	390.04	0.50	3,820.90
	25	163	80.40	11.52	219.52	0.27	2,193.10
	50	163	54.19	5.56	180.43	0.15	1,919.60
	75	163	31.22	3.32	87.61	0.07	710.20
	90	163	19.98	2.00	52.65	0.02	345.70
	95	163	15.80	1.17	44.33	0.01	323.30

Table 2.3. Summary statistics of candidate independent variables used for multiple linear regression models of a series of exceedance frequencies for annual and four seasonal time windows.

Independent variable	n	Mean	Median	SD	Min	Max
Drainage area; km^2	163	3080.42	487.54	5654.57	50.23	23316.70
Mean annual precipitation; mm	163	1319.17	1323.47	98.09	1072.78	1588.31
Mean annual temperature; °C	163	11.97	12.01	1.19	8.19	14.42
Mean annual humidity; %	163	69.44	69.16	1.86	65.01	72.68
Catchment slope; %	163	0.3066	0.3157	0.0770	0.0920	0.5424
Channel slope	163	0.207988	0.158000	0.176338	0.006000	0.964000
Site elevation; m	163	296.4	267.3	166.3	44.6	782.6
Urban; proportion	163	0.07439	0.03908	0.09512	0.00000	0.48800
Agriculture; proportion	163	0.23	0.20	0.13	0.00	0.85
Forest; proportion	163	0.63	0.65	0.18	0.00	1.00
Soil penetration rate 1; proportion	163	0.47412	0.48763	0.18100	0.00000	1.00000
Soil penetration rate 2; proportion	163	0.08623	0.04323	0.11319	0.00000	0.69675
Soil penetration rate 6; proportion	163	0.00061	0.00000	0.00313	0.00000	0.03564
Number of dams	163	1.48	0.00	2.67	0.00	11.00

Table 2.4. Summary statistics of 5 and 95 percent exceedance discharges for each major watershed. Stream discharge is represented by “Q” and “n” indicates the total number of gages for the summary of stream discharges and “SD” indicates standard deviation.

Dependent	Time	Percent	n	Mean	Median	SD	Min	Max
Geum River	Annual	05	38	141.55	49.35	236.85	4.75	1,051.50
		95	38	7.38	0.51	24.05	0.01	142.10
	January	05	38	77.95	12.46	173.45	0.53	814.99
		95	38	9.06	0.49	29.49	0.01	169.93
	July	05	38	379.08	158.32	563.50	33.00	2,379.30
		95	38	6.54	1.29	11.29	0.06	42.50
	May	05	38	106.92	28.27	215.54	2.73	1,104.40
		95	38	9.62	0.78	32.90	0.01	199.20
	October	05	38	91.63	17.83	195.30	0.50	933.00
		95	38	6.84	0.80	26.16	0.01	159.62
Han River	Annual	05	43	481.37	121.21	978.92	10.72	5,253.50
		95	43	22.81	1.23	55.39	0.08	243.00
	January	05	43	293.04	42.57	873.09	0.39	5,253.50
		95	43	17.28	1.00	47.29	0.00	243.00
	July	05	43	1,167.69	300.26	2,081.94	46.58	7,730.90
		95	43	35.07	3.37	83.21	0.17	415.10
	May	05	43	289.32	92.04	488.23	4.57	2,357.40
		95	43	24.74	2.00	60.07	0.07	243.00
	October	05	43	278.41	85.61	527.97	2.56	2,357.40
		95	43	22.09	1.29	54.76	0.08	243.01
Nakdong River	Annual	05	47	517.85	186.57	742.58	18.73	3,824.80
		95	47	21.31	2.58	49.42	0.03	285.70
	January	05	47	204.90	54.95	575.14	0.66	3,829.60
		95	47	16.47	1.51	51.05	0.02	341.80
	July	05	47	1,243.32	329.00	1,573.89	53.88	5,577.90
		95	47	40.93	5.99	73.83	0.04	357.80
	May	05	47	339.49	95.70	605.99	7.00	3,823.80
		95	47	20.73	4.00	48.39	0.03	272.60
	October	05	47	234.88	77.13	580.28	2.68	3,820.90
		95	47	26.21	2.62	56.96	0.05	323.30
Youngsum River	Annual	05	35	102.47	52.99	144.75	4.35	796.97
		95	35	3.03	0.75	7.52	0.01	43.32
	January	05	35	40.96	14.17	124.67	0.58	747.01
		95	35	2.54	0.55	5.65	0.01	31.49
	July	05	35	279.31	193.94	273.57	10.34	1,012.60
		95	35	5.29	1.78	10.64	0.06	60.61
	May	05	35	74.67	36.65	167.99	2.68	1,011.90
		95	35	2.93	0.65	8.91	0.03	52.40
	October	05	35	51.86	19.24	104.58	2.67	607.30
		95	35	3.85	0.97	9.30	0.02	53.31

Table 2.5. Multiple linear regression models of all exceedance flows for annual and seasonal time periods. Bold indicates significance at p ≤ 0.05, and bold and italics indicate significance at p ≤ 0.01.

Dependent variable	lnQ05	lnQ10	lnQ25	lnQ50	lnQ75	lnQ90	lnQ95
				Annual			
R^2	83.0	77.9	74.1	82.1	82.7	80.7	77.0
Constant	**-10.6123**	**-16.3673**	***-19.4971***	***-18.4329***	***-19.2821***	***-19.4391***	**-16.0484**
ln(drainage area); km^2	***0.831106***	***0.864626***	***0.859729***	***0.780459***	***0.689607***	***0.560106***	***0.443429***
ln(mean annual precipitation);	1.32675	**2.00645**	**2.47032**	***2.71222***	***2.92551***	***2.98887***	***2.53048***
ln(urban); proportion			***0.268396***	***0.251722***	***0.224077***	***0.194195***	**0.17517**
ln(soil penetration rate 1);	**-0.219328**	**-0.283668**					
ln(soil penetration rate 6);				**0.702762**	***0.870402***	***0.872046***	**0.800399**
number of dams				**0.0925863**	***0.146049***	***0.199773***	***0.241455***
				January			
R^2	58.4	64.3	67.7	78.8	78.7	61.1	55.5
Constant	**-23.0464**	**-18.1362**	**-19.9861**	**-16.1886**	**-13.3411**	***-2.88507***	***-2.78624***
ln(drainage area); km^2	***0.826205***	***0.648667***	***0.669642***	***0.687935***	***0.606028***	**0.51548**	***0.504523***
ln(mean annual precipitation);	**3.03713**	**2.46888**	**2.62569**	***2.56432***	***2.31494***		
ln(urban); proportion	***0.313385***	***0.353501***	***0.306212***	***0.240436***	***0.238819***		
ln(soil penetration rate 2);						***-0.19181***	-0.138359
ln(soil penetration rate 6);				**0.929796**	***1.14009***		
number of dams		**0.133555**	**0.120096**	***0.143135***	***0.165306***	**0.136834**	0.0965312

Table 2.5. Continued.

Dependent variable	lnQ05	lnQ10	lnQ25	lnQ50	lnQ75	lnQ90	lnQ95
				July			
R^2	89.9	89.9	86.4	88.3	88.5	85.5	82.4
Constant	***-13.4167***	***-15.2878***	***-19.2261***	***-21.939***	***-22.7659***	***-18.5867***	***-20.6566***
ln(drainage area); km^2	***0.804739***	***0.840276***	***0.880323***	***0.898989***	***0.802098***	***0.680423***	***0.586347***
ln(mean annual precipitation);	***1.89231***	***2.04664***	***2.52389***	***2.5777***	***3.07318***	***2.68226***	***2.59887***
ln(catchment slope); %				**0.457961**	**0.423582**		
ln(urban); proportion			***0.153663***	***0.257814***	***0.269509***	***0.187887***	**0.169556**
ln(soil penetration rate 1);	***-0.19048***	***-0.204683***					
ln(soil penetration rate 6);					**0.627247**	**0.637503**	
number of dams					***0.0895241***	***0.166023***	***0.208082***
				May			
R^2	77.4	75.2	71.8	73.5	78.7	76.3	73.9
Constant	**-16.3871**	***-21.8461***	***-22.544***	***-22.8615***	***-20.9683***	**-15.6768**	**-15.2951**
ln(drainage area); km^2	***0.846564***	***0.866198***	***0.862279***	***0.889255***	***0.72048***	***0.5627***	***0.480752***
ln(mean annual precipitation);	**2.14646**	***2.85731***	***2.89372***	***2.83818***	***3.20758***	***2.57554***	***2.43641***
ln(urban); proportion	**0.173597**	***0.228883***	***0.277603***	**0.323221**	***0.249392***	***0.232406***	***0.216896***
ln(soil penetration rate 6);					**0.953539**	***1.00974***	**0.821956**
number of dams					***0.129692***	***0.180829***	***0.209329***
				October			
R^2	65.3	67.1	71.3	81.9	81.5	81.6	78.8
Constant	***-25.7***	**-21.1433**	***-23.291***	***-23.0566***	***-21.1528***	***-17.4865***	**-12.9097**
ln(drainage area); km^2	***0.837724***	***0.829654***	***0.860611***	***0.78783***	***0.692558***	***0.587187***	***0.465319***
ln(mean annual precipitation);	***3.44102***	**2.77036**	***2.9632***	***2.87891***	***2.63358***	***2.14869***	1.56055
ln(urban); proportion	0.280296	***0.274625***	***0.270554***	***0.224122***	***0.230559***	***0.189222***	0.126773
number of dams				**0.100916**	***0.150598***	***0.195152***	***0.24935***

Table 2.6. Table 6. Mean absolute percentage errors (MAEs) and Nash-Sutcliffe model efficiency coefficient (NSE) of exceedance flow models for annual and four seasonal time windows.

Time windows	Percent exceedances	Generation step		Validation step	
		MAE [a]	NSE [b]	MAE [a]	NSE [b]
Annual	Q05	42.18	0.83	39.19	0.83
	Q10	51.95	0.78	49.70	0.73
	Q25	58.73	0.74	59.21	0.61
	Q50	47.33	0.82	56.23	0.75
	Q75	44.00	0.83	43.42	0.82
	Q90	46.53	0.81	40.28	0.84
	Q95	46.85	0.77	38.71	0.85
	Mean	**48.22**	**0.80**	**46.68**	**0.78**
January	Q05	85.57	0.58	95.23	0.57
	Q10	74.47	0.64	91.86	0.47
	Q25	64.29	0.68	86.19	0.50
	Q50	49.80	0.79	57.62	0.75
	Q75	49.39	0.79	41.94	0.84
	Q90	62.82	0.61	40.56	0.85
	Q95	63.71	0.56	43.28	0.81
	Mean	**64.29**	**0.66**	**65.24**	**0.68**
July	Q05	29.22	0.90	33.27	0.90
	Q10	31.53	0.90	36.81	0.85
	Q25	38.19	0.86	40.74	0.80
	Q50	38.09	0.88	43.16	0.83
	Q75	38.25	0.88	43.29	0.84
	Q90	40.85	0.85	46.27	0.81
	Q95	43.95	0.82	52.61	0.77
	Mean	**37.15**	**0.87**	**42.31**	**0.83**
May	Q05	53.47	0.77	51.00	0.65
	Q10	58.15	0.75	54.22	0.65
	Q25	60.78	0.72	57.47	0.65
	Q50	58.66	0.74	61.30	0.67
	Q75	52.45	0.79	52.15	0.76
	Q90	54.11	0.76	50.55	0.77
	Q95	54.12	0.74	51.05	0.72
	Mean	**55.96**	**0.75**	**53.96**	**0.70**
October	Q05	73.18	0.65	69.21	0.68
	Q10	72.17	0.67	67.41	0.66
	Q25	61.85	0.71	67.45	0.64
	Q50	47.16	0.82	50.20	0.76
	Q75	45.89	0.81	39.36	0.82
	Q90	43.58	0.82	43.03	0.79
	Q95	46.63	0.79	47.15	0.70
	Mean	**55.78**	**0.75**	**54.83**	**0.72**

Table 2.7. Two tailed Pearson correlation tests between observed and predicted stream discharges for all sites combined (n=163) and four major watersheds. Bold indicates significance at p≤0.05, and bold and italics indicate significance at p≤0.01.

	Annual	January	July	May	October	Mean
			All sites (n=163)			
lnQ05	***0.912***	***0.762***	***0.948***	***0.872***	***0.811***	0.861
lnQ10	***0.880***	***0.784***	***0.945***	***0.861***	***0.819***	0.858
lnQ25	***0.852***	***0.805***	***0.926***	***0.843***	***0.837***	0.853
lnQ50	***0.901***	***0.885***	***0.937***	***0.852***	***0.900***	0.895
lnQ75	***0.910***	***0.892***	***0.937***	***0.885***	***0.904***	0.906
lnQ90	***0.902***	***0.805***	***0.921***	***0.874***	***0.902***	0.881
lnQ95	***0.885***	***0.773***	***0.904***	***0.859***	***0.883***	0.861
Mean	0.892	0.815	0.931	0.864	0.865	
			Geum R. (n=38)			
lnQ05	***0.951***	***0.760***	***0.958***	***0.909***	***0.871***	0.890
lnQ10	***0.908***	***0.825***	***0.962***	***0.887***	***0.864***	0.889
lnQ25	***0.842***	***0.796***	***0.916***	***0.824***	***0.821***	0.840
lnQ50	***0.897***	***0.834***	***0.924***	***0.783***	***0.863***	0.860
lnQ75	***0.872***	***0.854***	***0.911***	***0.804***	***0.825***	0.853
lnQ90	***0.857***	***0.804***	***0.832***	***0.811***	***0.806***	0.822
lnQ95	***0.844***	***0.784***	***0.758***	***0.790***	***0.711***	0.777
Mean	0.882	0.808	0.894	0.830	0.823	
			Han R. (n=43)			
lnQ05	***0.925***	***0.776***	***0.964***	***0.853***	***0.809***	0.865
lnQ10	***0.893***	***0.750***	***0.964***	***0.851***	***0.830***	0.858
lnQ25	***0.869***	***0.804***	***0.945***	***0.835***	***0.850***	0.861
lnQ50	***0.913***	***0.892***	***0.951***	***0.866***	***0.918***	0.908
lnQ75	***0.919***	***0.895***	***0.946***	***0.900***	***0.924***	0.917
lnQ90	***0.932***	***0.764***	***0.945***	***0.893***	***0.929***	0.893
lnQ95	***0.921***	***0.766***	***0.937***	***0.882***	***0.901***	0.881
Mean	0.910	0.807	0.950	0.869	0.880	
			Nakdong R. (n=47)			
lnQ05	***0.901***	***0.802***	***0.963***	***0.905***	***0.833***	0.881
lnQ10	***0.882***	***0.817***	***0.954***	***0.895***	***0.840***	0.878
lnQ25	***0.899***	***0.865***	***0.945***	***0.919***	***0.889***	0.903
lnQ50	***0.909***	***0.924***	***0.941***	***0.884***	***0.915***	0.915
lnQ75	***0.925***	***0.903***	***0.947***	***0.917***	***0.925***	0.923
lnQ90	***0.887***	***0.823***	***0.931***	***0.901***	***0.905***	0.889
lnQ95	***0.860***	***0.750***	***0.916***	***0.879***	***0.902***	0.861
Mean	0.895	0.841	0.942	0.900	0.887	
			Youngsum R. (n=35)			
lnQ05	***0.774***	***0.445***	***0.834***	***0.663***	***0.520***	0.647
lnQ10	***0.709***	***0.510***	***0.816***	***0.625***	***0.510***	0.634
lnQ25	***0.595***	***0.510***	***0.780***	***0.574***	***0.581***	0.608
lnQ50	***0.755***	***0.732***	***0.849***	***0.713***	***0.781***	0.766
lnQ75	***0.810***	***0.841***	***0.860***	***0.777***	***0.810***	0.820
lnQ90	***0.844***	***0.789***	***0.850***	***0.724***	***0.846***	0.811
lnQ95	***0.837***	***0.782***	***0.829***	***0.688***	***0.831***	0.793
Mean	0.761	0.658	0.831	0.681	0.697	

Table 2.8. Principal components analysis (PCA). Stream flow characteristics variable loadings for PC 1, 2, and 3 (n= 163). Bold values are considered equal or larger than |0.40|, but PC3 was not considered because its eigen value (0.190) was less than 1.

Parameter	PC1	PC2	PC3
Stream power (Q10×site slope/flow width)	-0.073	**-0.980**	-0.186
Baseflow yield (Q90/catchment area)	**-0.710**	-0.079	0.699
Flow flashiness (Q10/Q90)	**0.700**	-0.183	0.690
Eigen value	1.786	1.025	0.190
Proportion of variance	59.5	34.2	6.3
Cumulative proportion	59.5	93.7	100

Table 2.9. Comparison of proportion (site numbers) of stream flow types by flow stability and stream power for stream flow gages and biological sampling sites.

Watersheds	Stream flow gages (n=163)						Biological sampling sites (n=684)					
	F	S	Total	HP	LP	Total	F	S	Total	HP	LP	Total
Overall	50.3% (82)	49.7% (81)	100.0% (163)	49.7% (81)	50.3% (82)	100.0% (163)	64.8% (443)	35.2% (241)	100.0% (684)	41.5% (284)	58.5% (400)	100.0% (684)
Geum R.	52.6% (20)	47.4% (18)	100.0% (38)	39.5% (15)	60.5% (23)	100.0% (38)	65.1% (80)	38.5% (50)	100.0% (130)	27.7% (36)	72.3% (94)	100.0% (130)
Han R.	44.2% (19)	55.8% (24)	100.0% (43)	62.8% (27)	37.2% (16)	100.0% (43)	65.1% (185)	34.9% (99)	100.0% (284)	46.8% (133)	53.2% (151)	100.0% (284)
Nakdong R.	48.9% (23)	51.1% (24)	100.0% (47)	59.6% (28)	40.4% (19)	100.0% (47)	73.1% (95)	26.9% (35)	100.0% (130)	44.6% (58)	55.4% (72)	100.0% (130)
Yeongsum R.	57.1% (20)	42.9% (15)	100.0% (35)	31.4% (11)	68.6% (24)	100.0% (35)	59.3% (83)	40.7% (57)	100.0% (140)	40.7% (57)	59.3% (83)	100.0% (140)

Table 2.10. Comparison of proportion (site numbers) of each stream flow type for stream flow gages and biological sampling sites.

Watersheds	Stream flow gages (n=163)					Biological sampling sites (n=684)				
	F-HP	F-LP	S-HP	S-LP	Total	F-HP	F-LP	S-HP	S-LP	Total
Overall	26.4% (43)	23.9% (39)	23.3% (38)	26.4% (43)	100.0% (163)	25.6% (175)	39.2% (268)	15.9% (109)	19.3% (132)	100.0% (684)
Geum R.	18.4% (7)	34.2% (13)	21.1% (8)	26.3% (10)	100.0% (38)	15.4% (20)	46.2% (60)	12.3% (16)	26.2% (34)	100.0% (130)
Han R.	30.2% (13)	14.0% (6)	32.6% (14)	23.3% (10)	100.0% (43)	29.2% (83)	35.9% (102)	17.6% (50)	17.3% (49)	100.0% (284)
Nakdong R.	29.8% (14)	19.1% (9)	29.8% (14)	21.3% (10)	100.0% (47)	30.3% (39)	43.1% (56)	14.6% (19)	12.3% (16)	100.0% (130)
Yeongsum R.	25.7% (9)	31.4% (11)	5.7% (2)	37.1% (13)	100.0% (35)	23.6% (33)	35.7% (50)	17.1% (24)	23.6% (33)	100.0% (140)

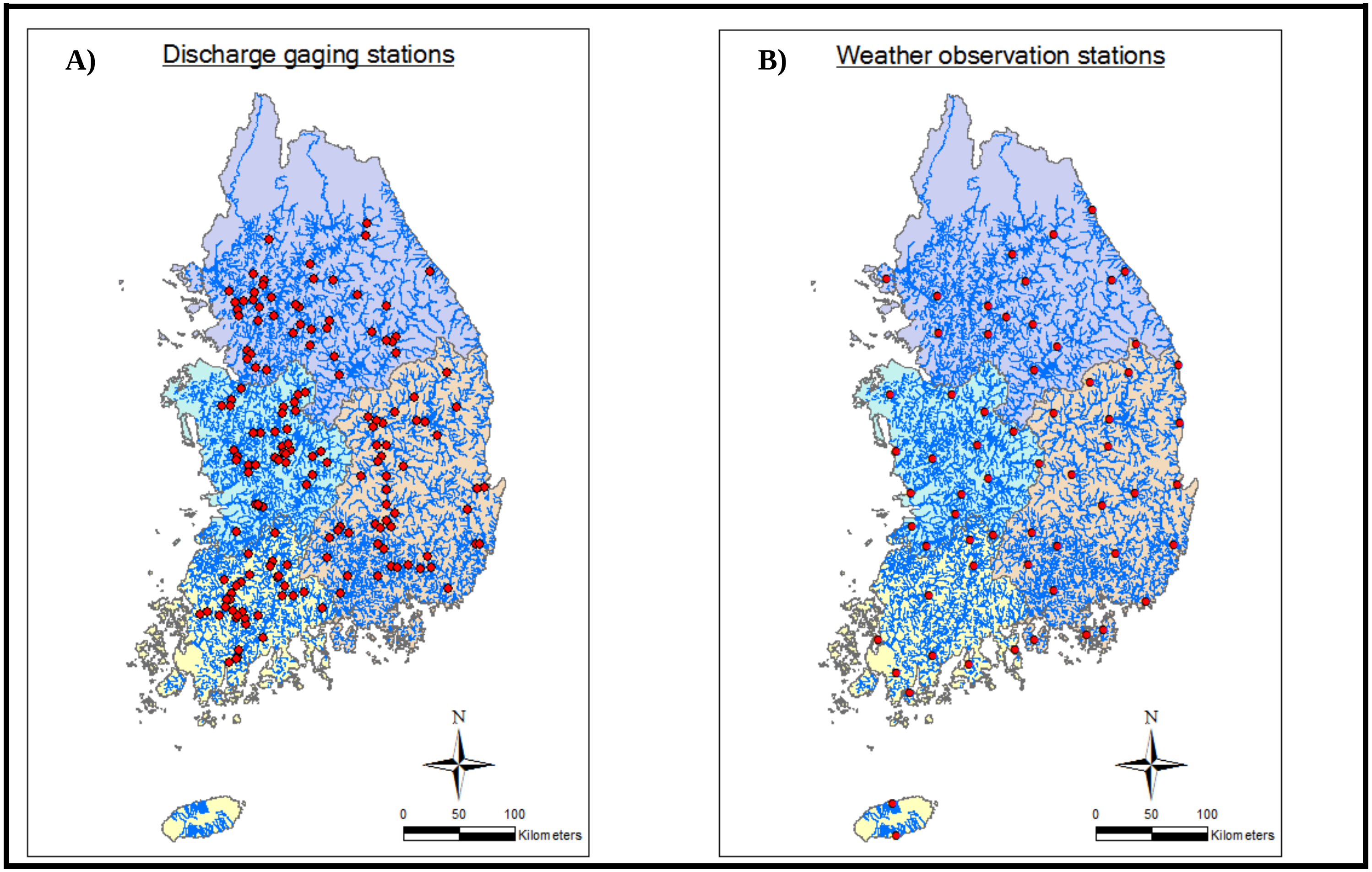


Figure 2.1. Locations of A) stream discharge gages (163 sites) and B) weather-observation stations (63 sites). Violet, sky blue, orange, and yellow colors indicate the Han River, Geum River, Nakdong River, and Youngsum River (Youngsan and Sumjin Rivers) Watersheds, respectively.

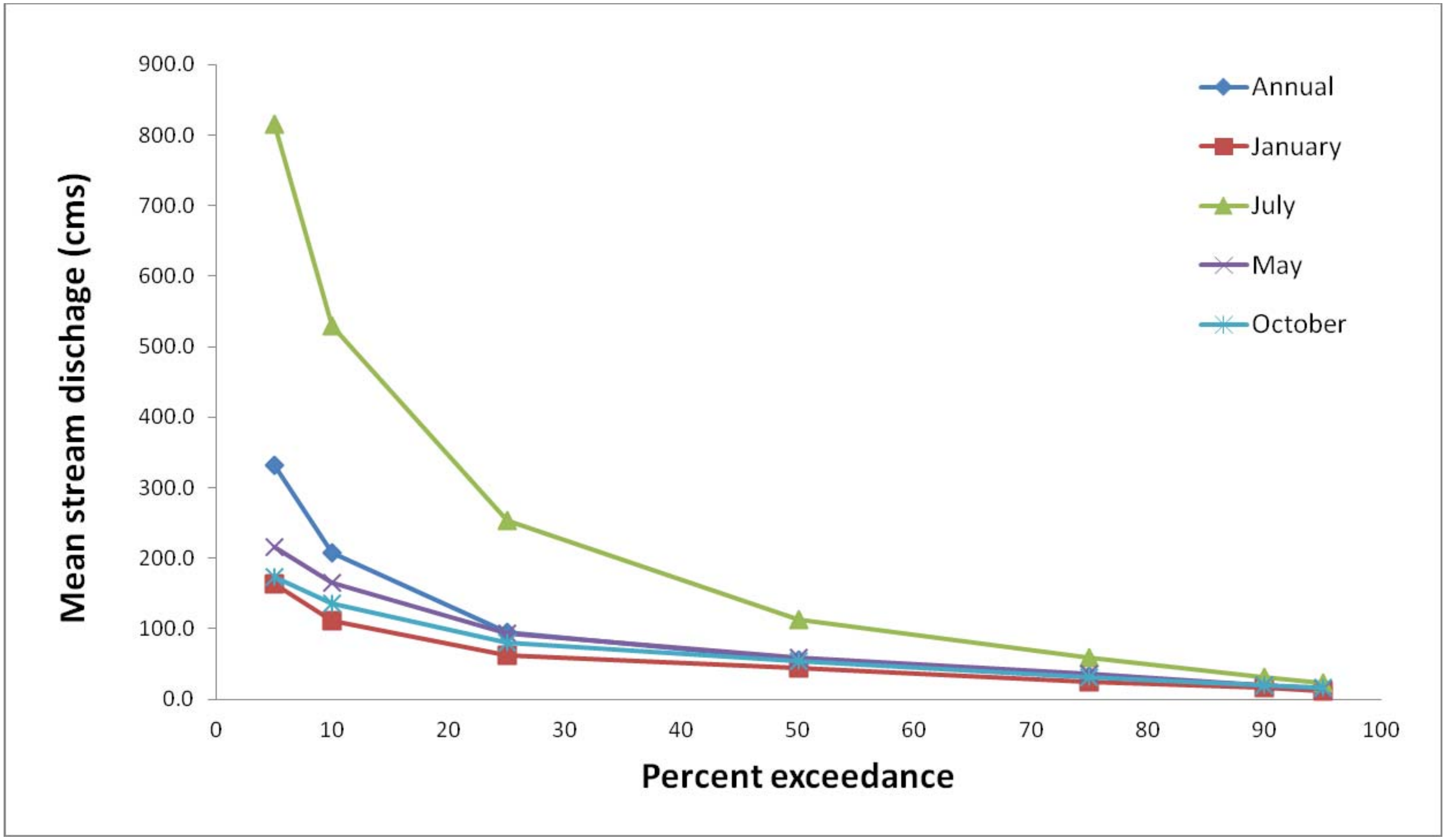


Figure 2.2. Comparison of mean stream discharges (cms) of percent exceedance freqencies for five time windows (Annual, January, May, July, and October) for all watersheds.

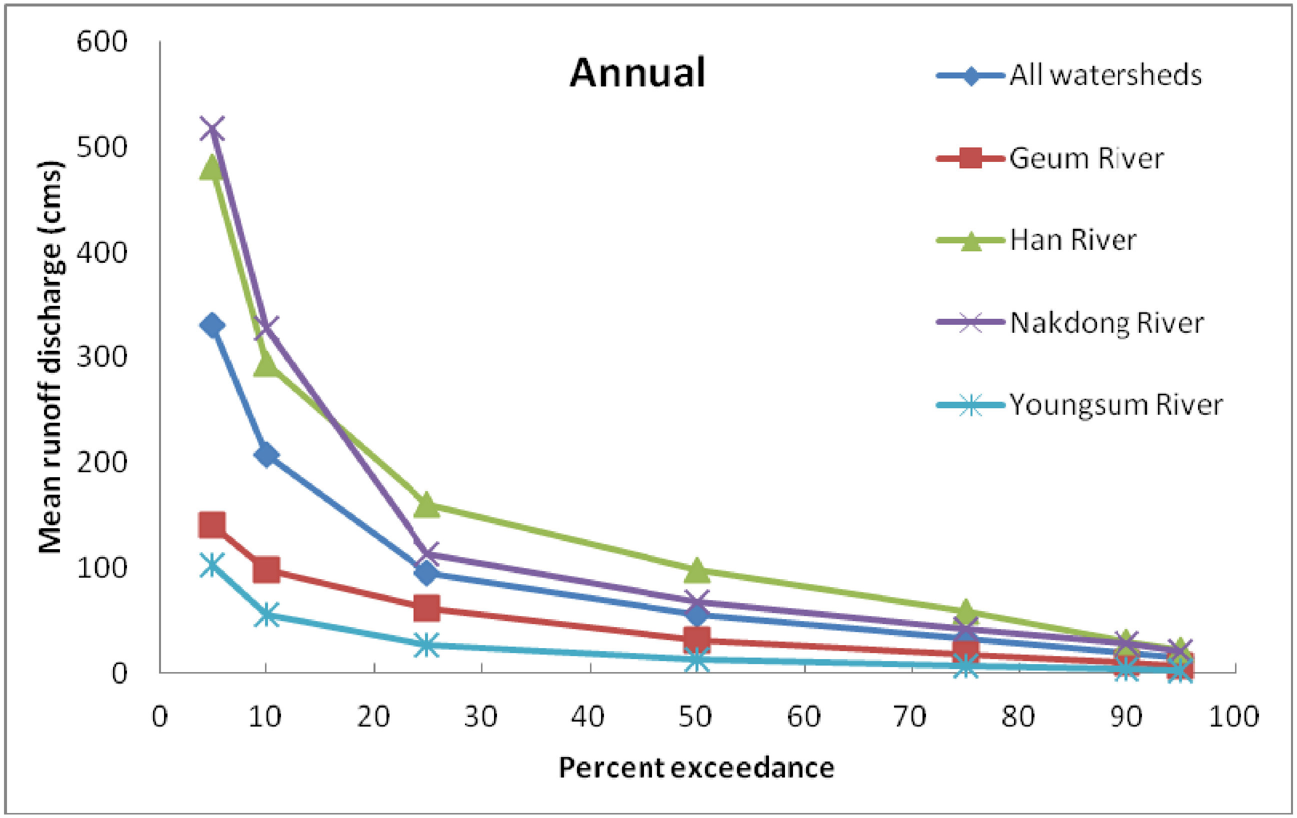


Figure 2.3. Comparison of mean stream discharges (cms) of each percent exceedance frequency for overall and four major Korean watersheds (Geum River, Han River, Nakdong River, and Youngsum (Youngsan and Sumjin) River).

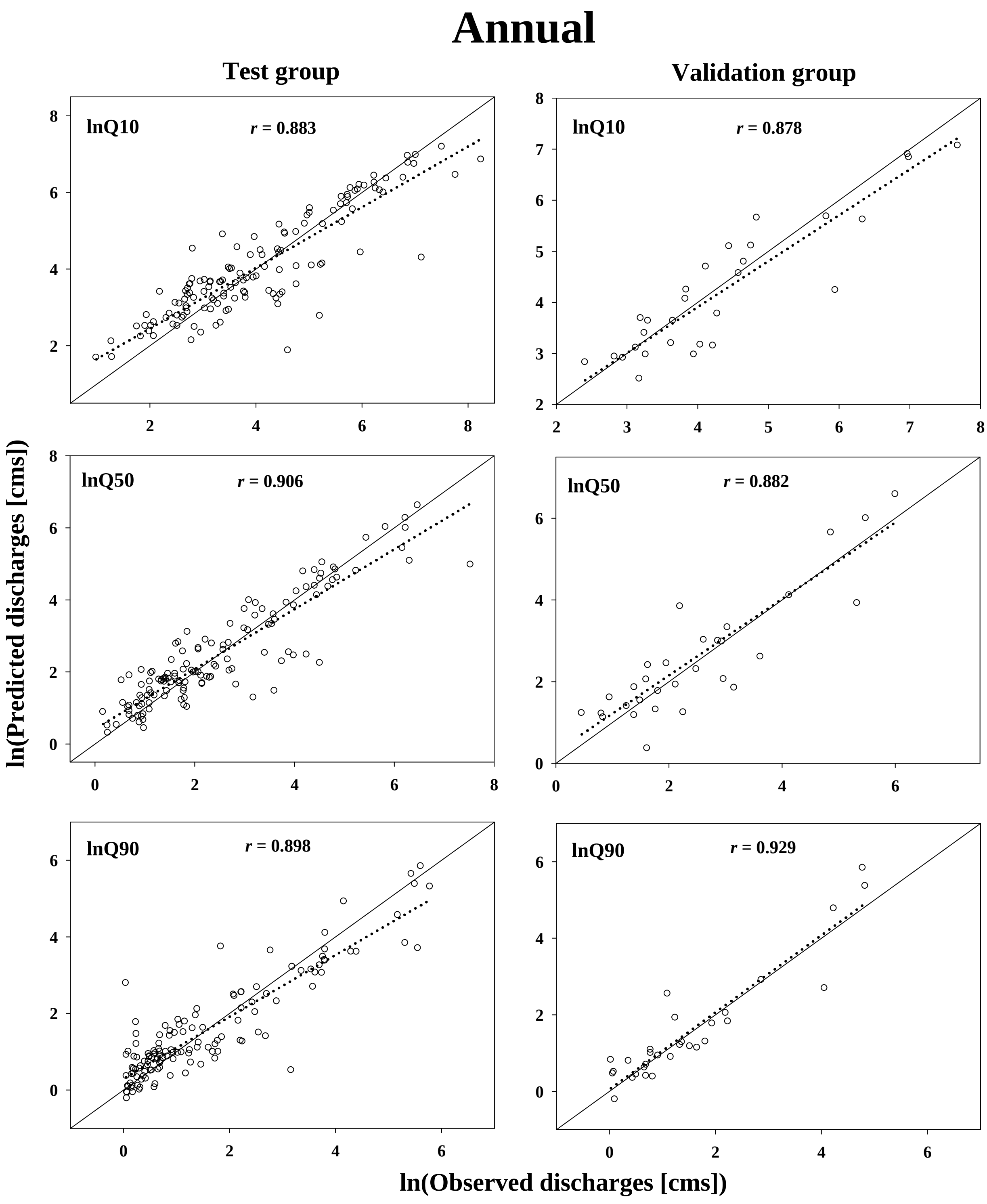


Figure 2.4. Scatter plots of the relationship between predicted and observed stream discharges for Annual time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line.

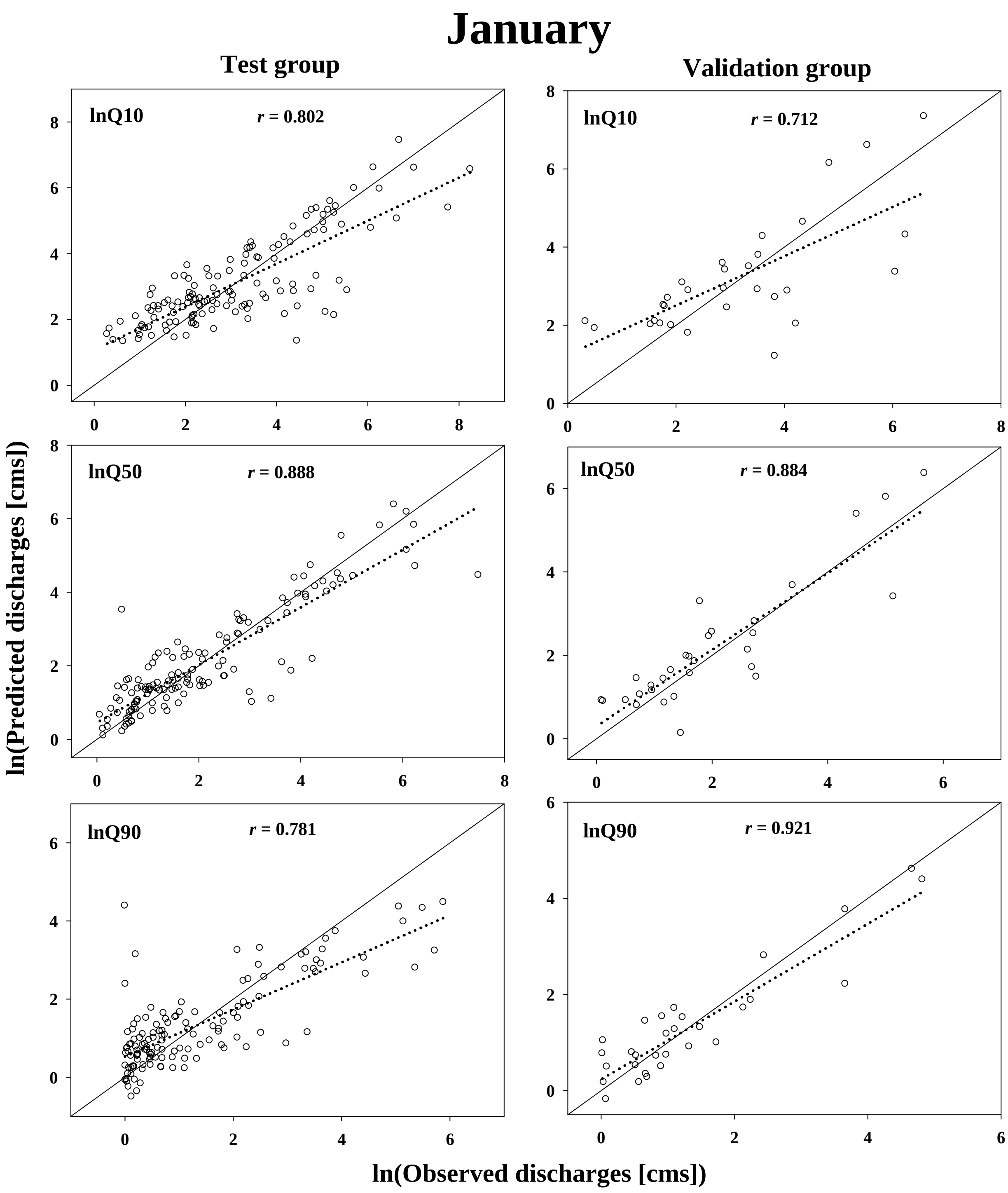


Figure 2.5. Scatter plots of the relationship between predicted and observed exceedance discharges for January time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line.

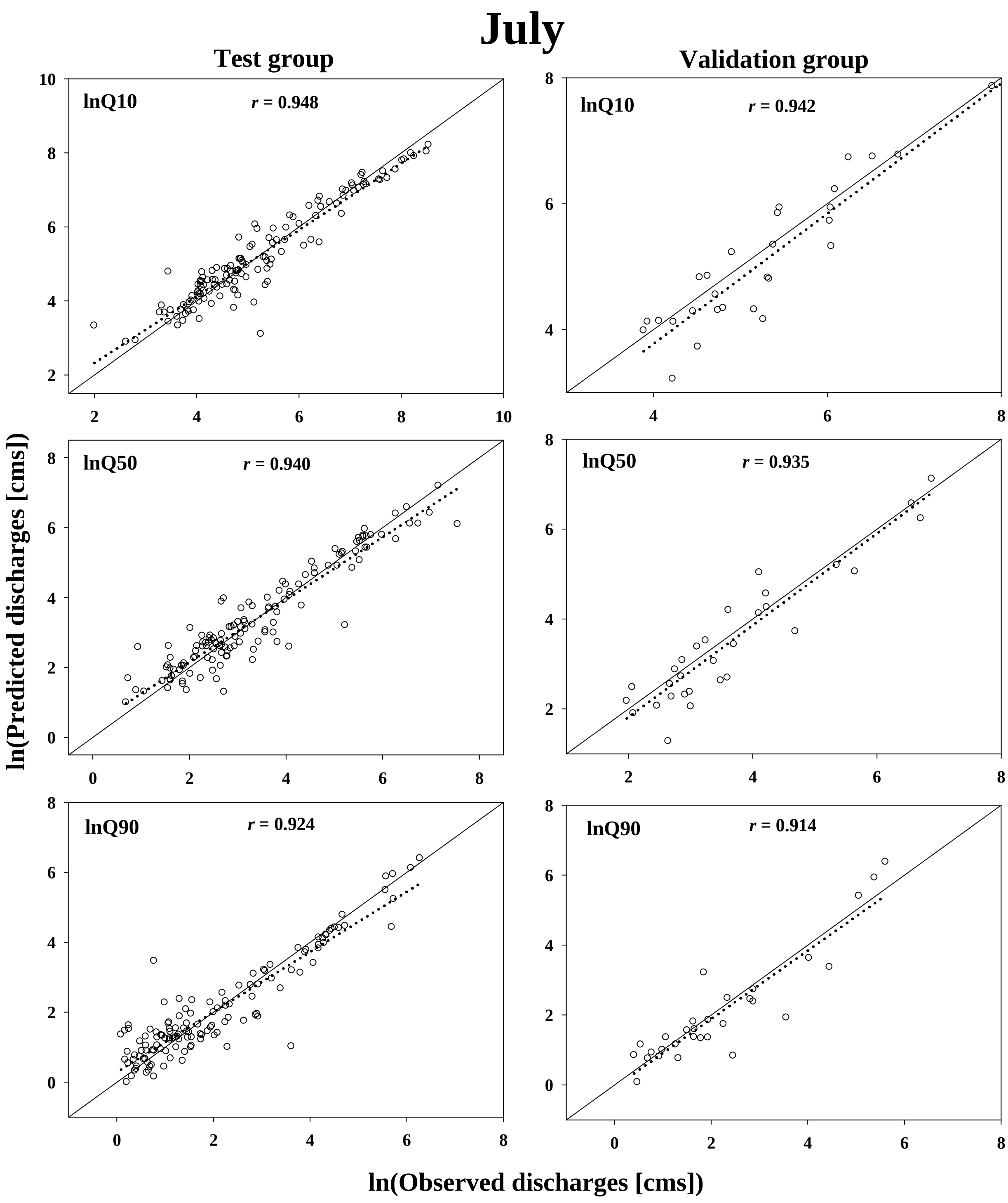


Figure 2.6. Scatter plots of the relationship between predicted and observed exceedance discharges for July time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line.

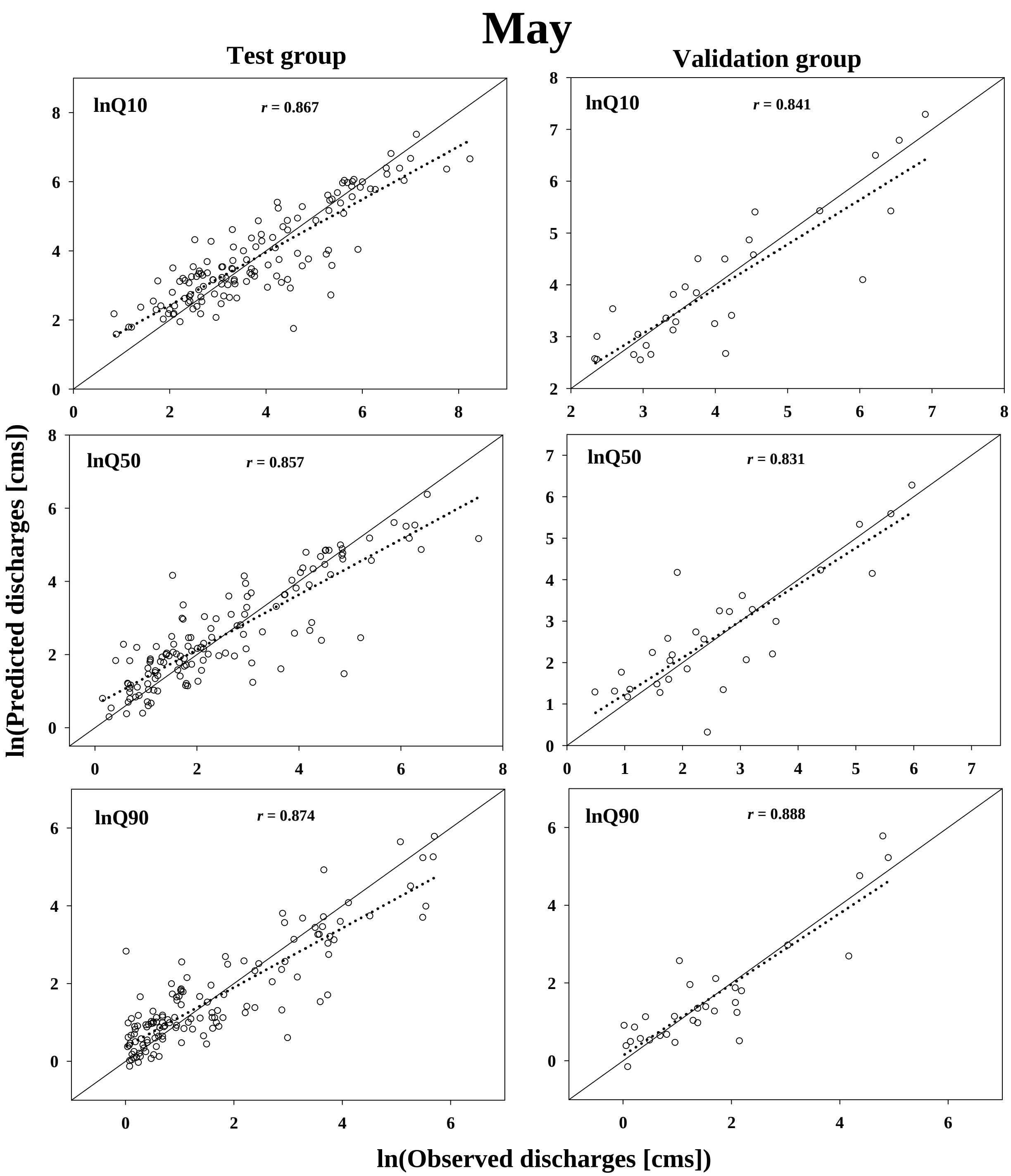


Figure 2.7. Scatter plots of the relationship between predicted and observed exceedance discharges for May time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line.

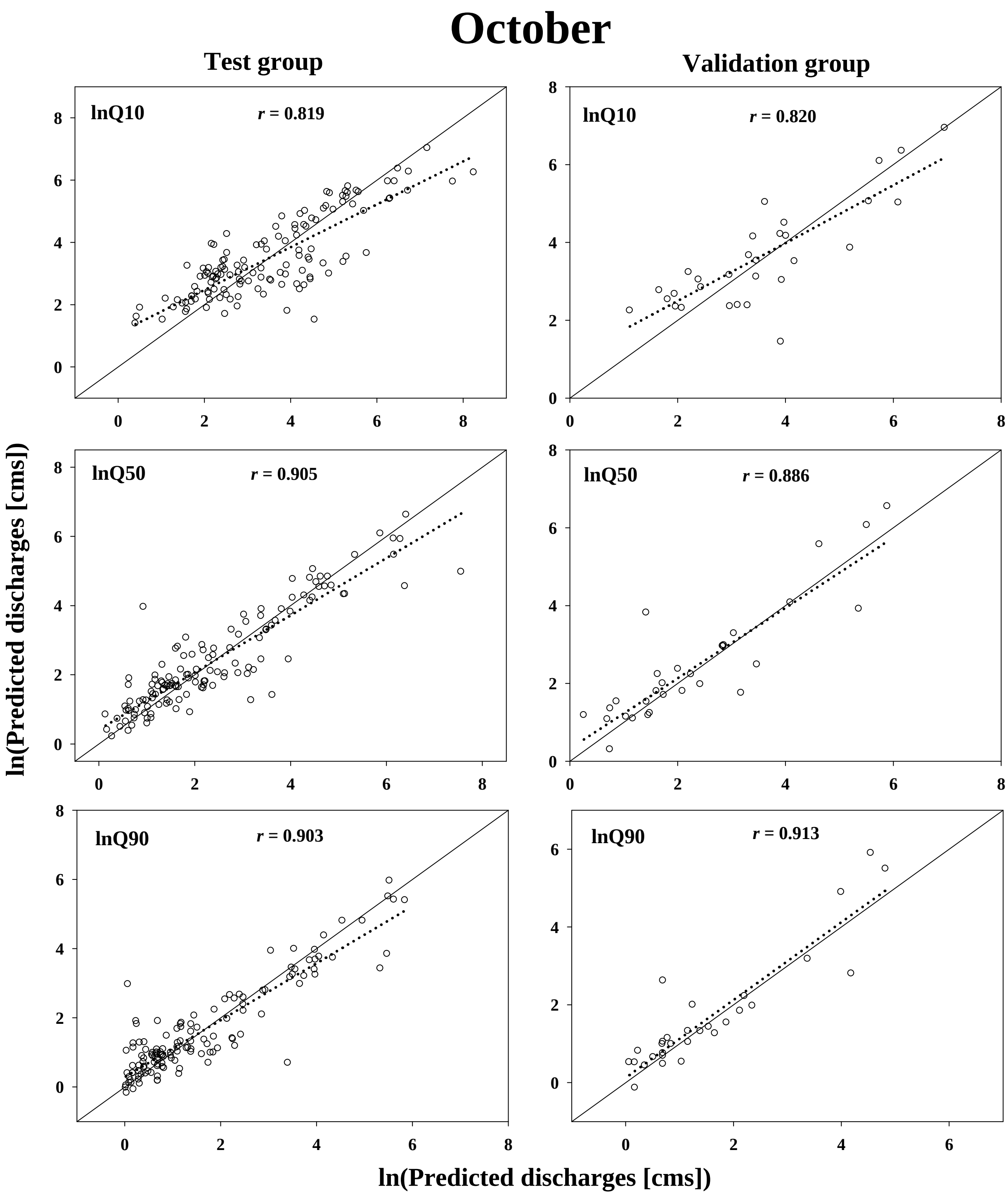


Figure 2.8. Scatter plots of the relationship between predicted and observed exceedance discharges for October time window. Left and right column panels show the test group (133 sites) and the validation group (30 sites), respectively. The ideal 1:1 relationship is shown as a solid line and the model relationship is shown as a dashed line.

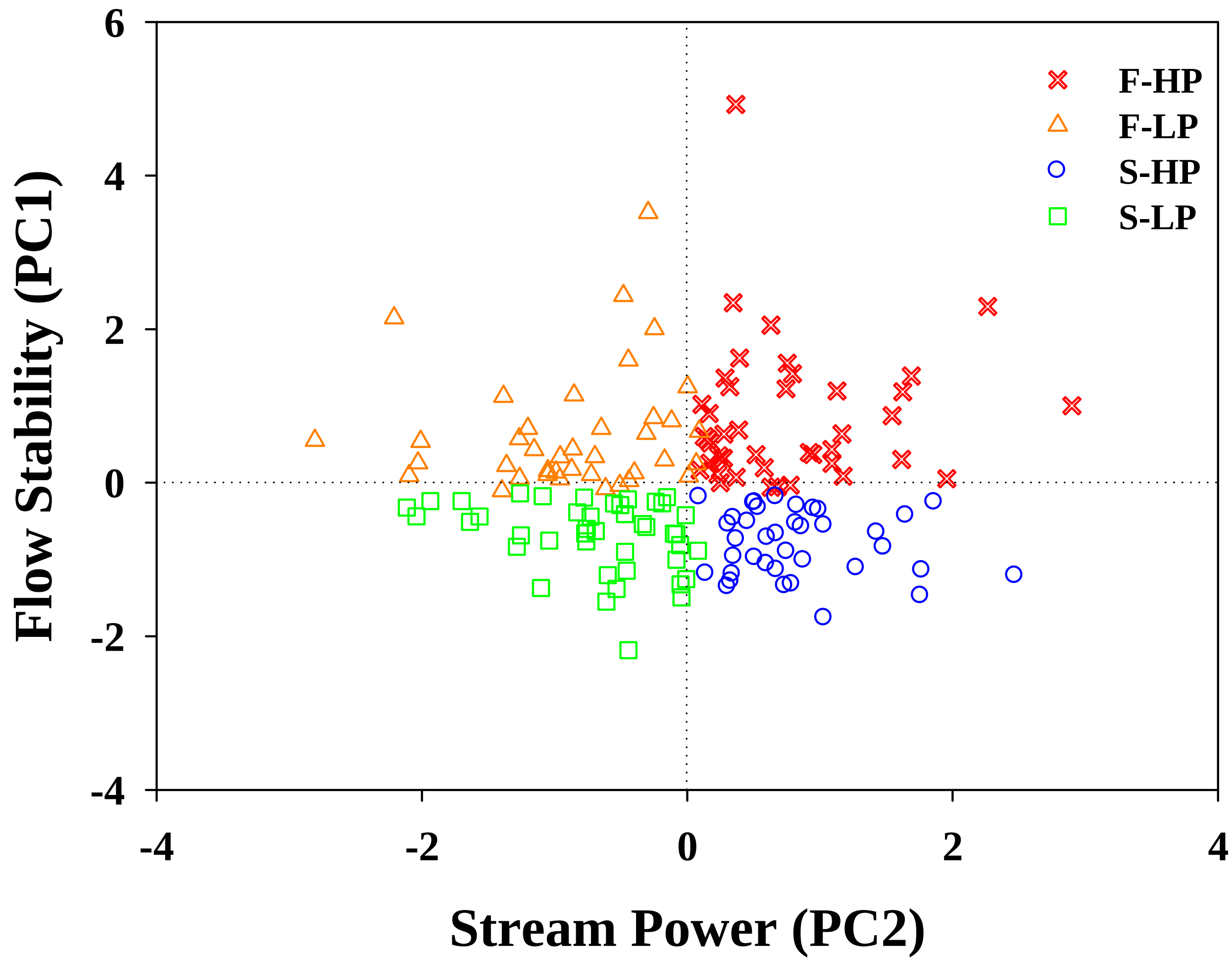


Figure 2.9. Two-dimensional PC score plot for the Korean stream classification using stream power and flow stability. Median numbers of both PC scores were used for stream classification. F, S, HP, and LP indicate flashier stream, more stable stream, higher-powered stream, and lower-powered stream, respectively.

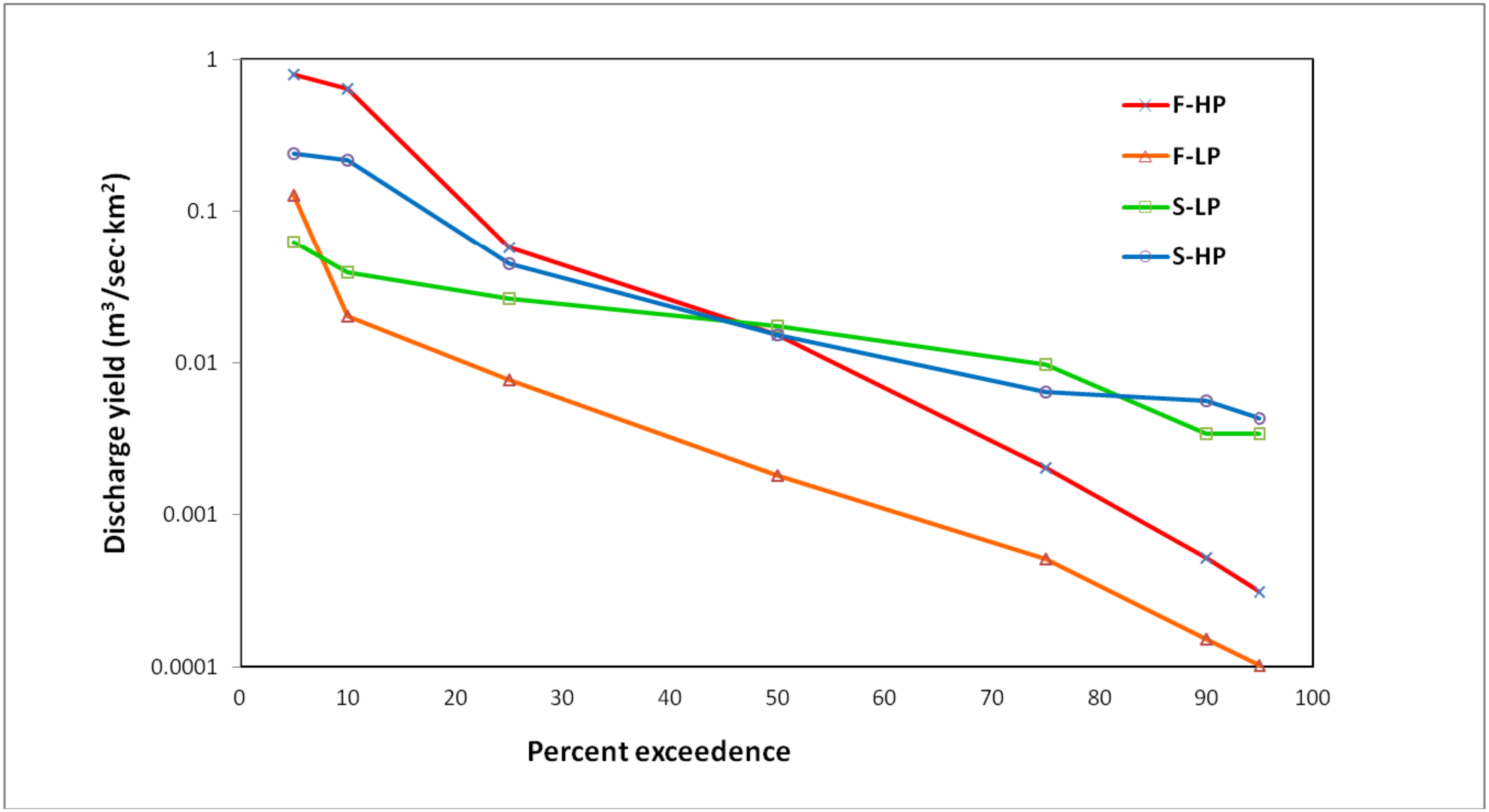


Figure 2.10. Example flow duration curves of four different stream flow types for Korean streams based on stream classification using PCA analysis. F, S, HP, and LP indicate flashier stream, more stable stream, higher-powered stream, and lower-powered stream, respectively.

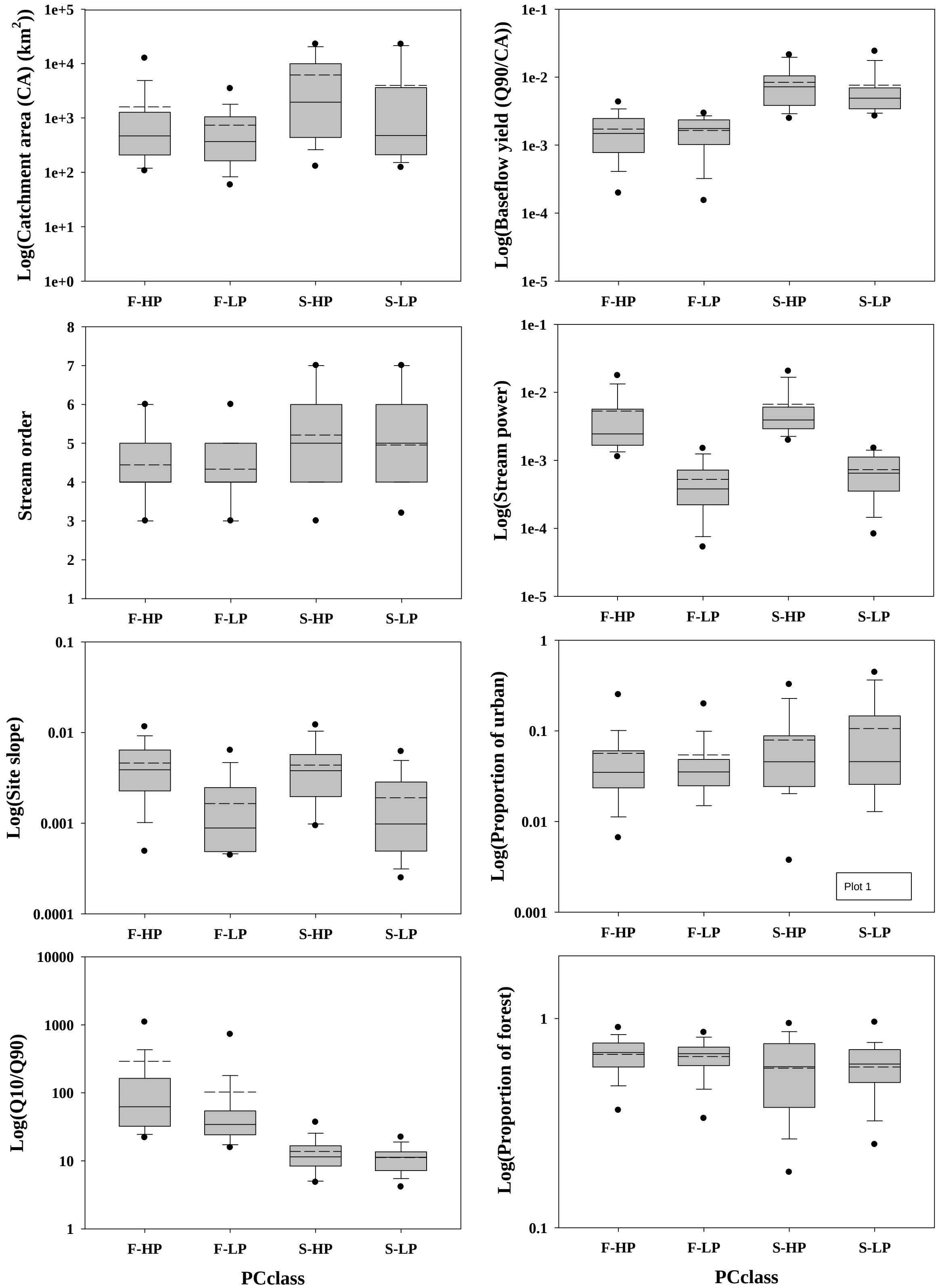


Figure 2.11. Comparison of landscape variables and stream flow characteristics among four different stream flow types for stream flow gages.

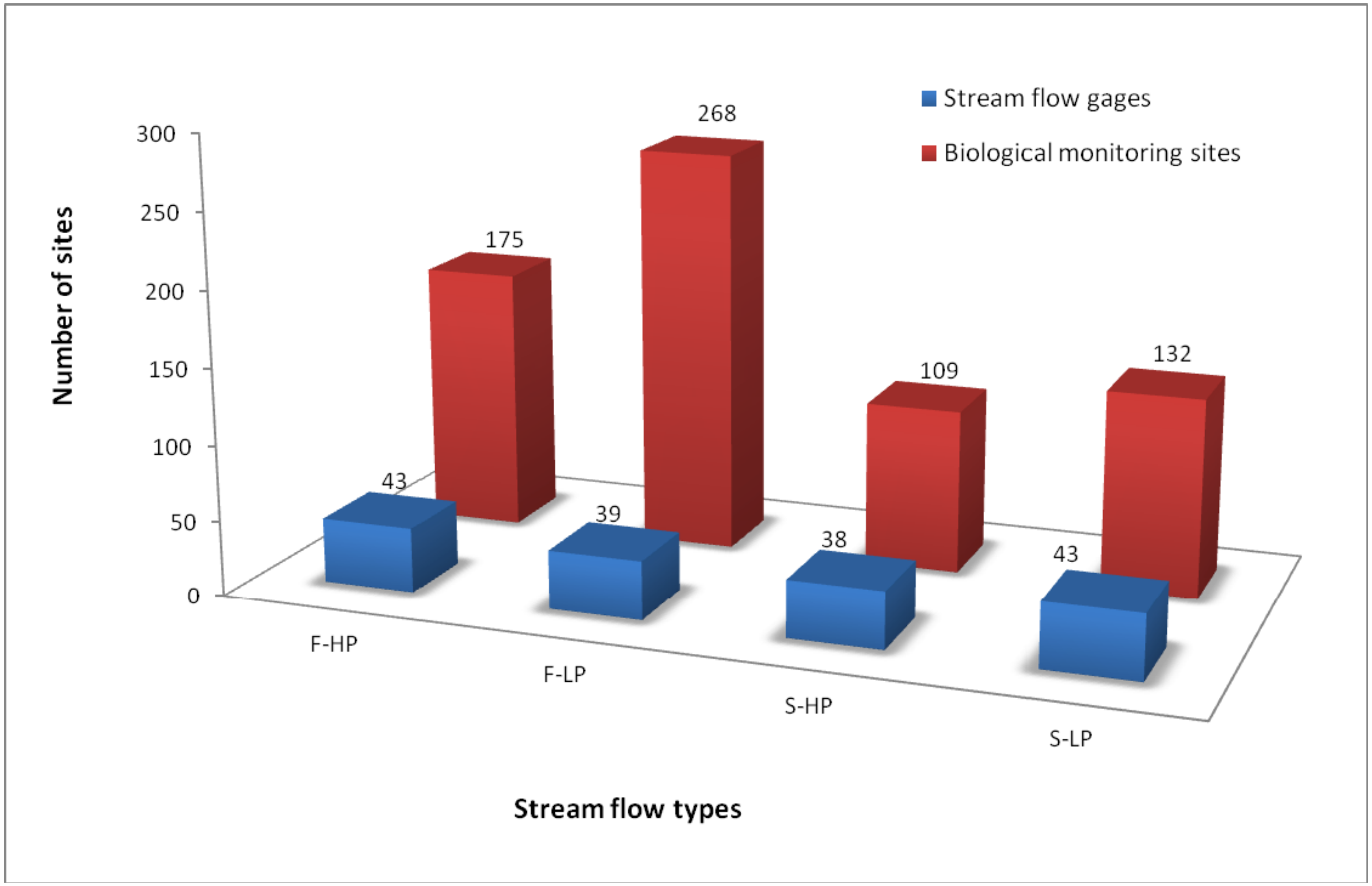
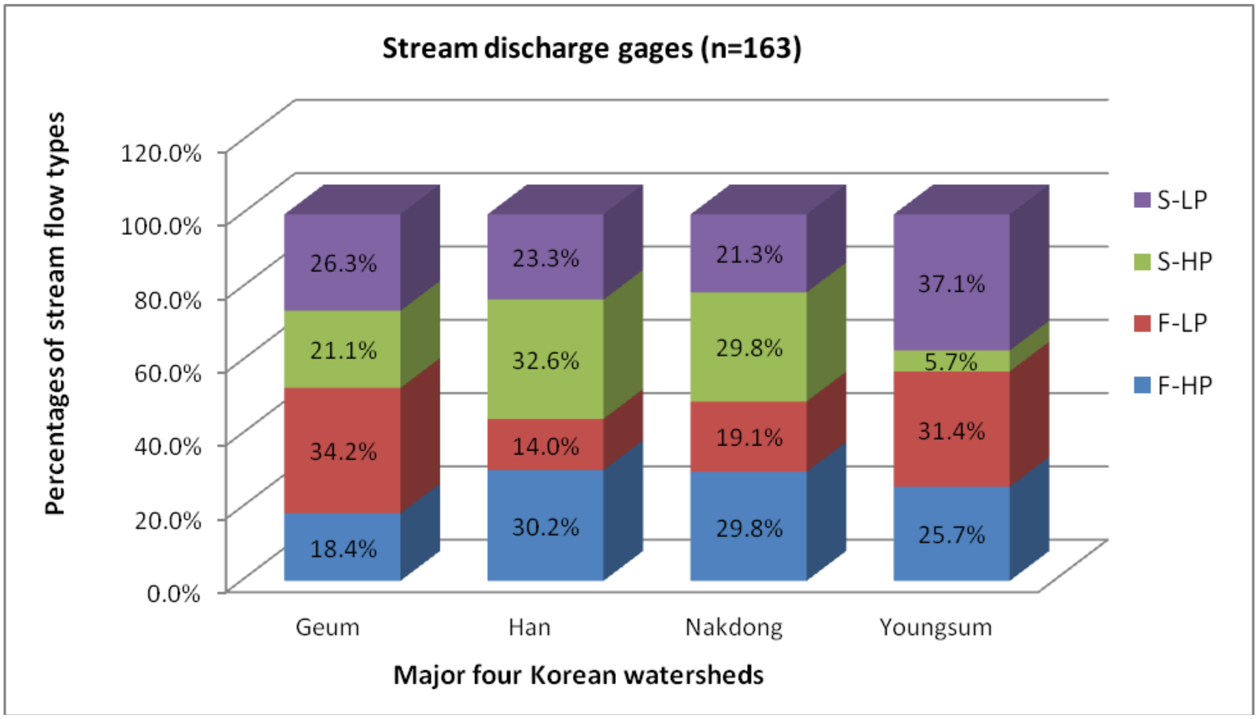


Figure 2.12. Comparison of site numbers for four different flow types between stream flow gages and biological monitoring sites. Stream flow type for each site was produced by PCA and PCR models. F, S, HP, and LP indicate flashier stream, more stable stream, higher power stream, and lower power stream, respectively.



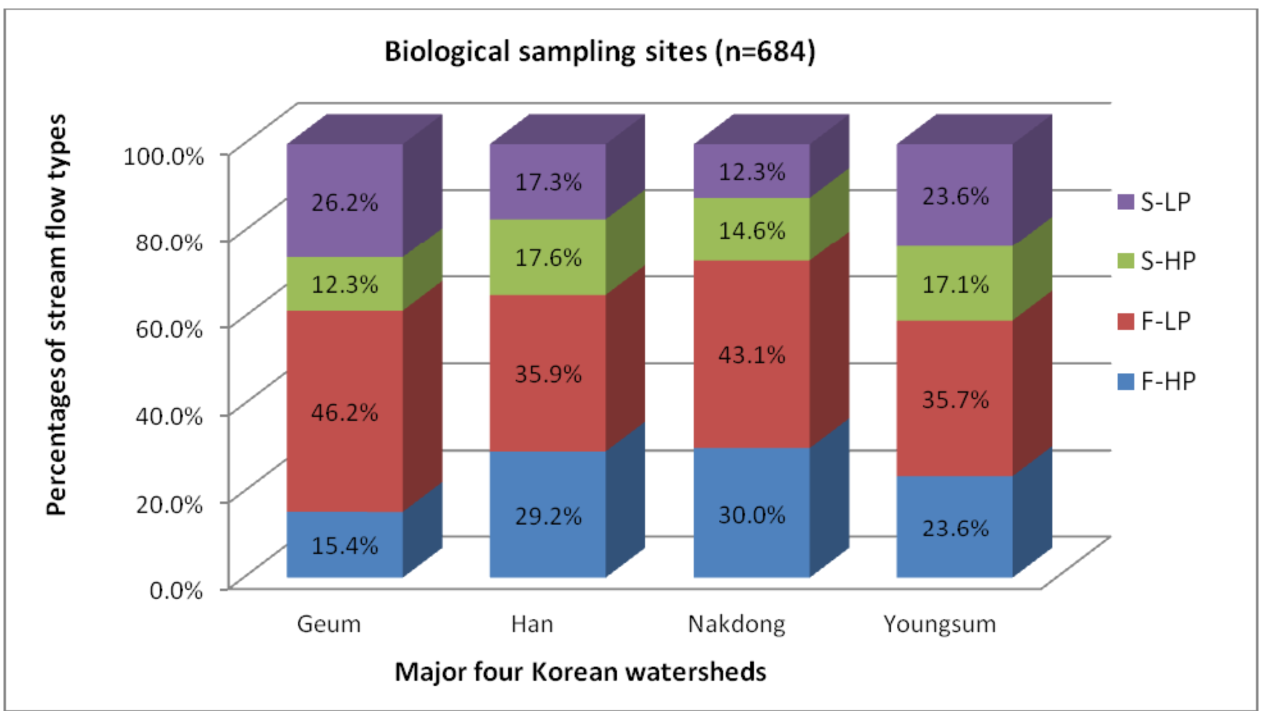


Figure 2.13. Comparison of percentages of each flow type for major four Korean watersheds. Top graph shows stream flow gages (n=163) and bottom graph shows biological monitoring sites (n=684 sites).

Literature cited

Allan, J. D. 2004. Landscapes and riverscapes: The influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics. 35: 257-284.

Allan, D. and L. Hinz. 2004. An assessment of flows for rivers of the Great Lakes Basin. Final report to the Great Lakes Protection Fund, Chicago, Illinois.

Allan, J. D. and M. M. Castillo. 2007. Stream ecology: structure and function of running waters. Second edition. Springer. Dordrecht, The Netherlands.

Anonymous. 2002. The National Map - Hydrography. U.S. Geological Survey. Fact Sheet 060-02. Reston, Virginia.

Anonymous. 2003. Great Lakes regional aquatic GAP analysis - Preserving biodiversity in the Great Lakes Basin. U.S. Geological Survey. Great Lakes Science Center. Fact Sheet 2003-1. Ann Arbor, Michigan.

Anonymous. 2008. Ecological limits of hydrologic alteration: environmental flows for regional water management. ELOHA Working Group. Fact Sheet. Available: http://www.nature.org/ELOHA. (May 2009).

Arthington, A. H., S. E. Bunn, N. L. Poff, and R. J. Naiman. 2006. The challenge of providing environmental flow rules to sustain river ecosystems. Ecological Applications. 16: 1311-1318.

Bae, M., Y. Kwon, S. Hwang, T. Chon, H. Yang, I. Kwak, J. Park, S. Ham, and Y. Park. 2011. Relationship between three major stream assemblages and their environmental factors in multiple spatial scales. International Journal of Limnology. 47: S91-S105.

Bagnold, R. A. 1966. An approach to the sediment transport problem from general physics. Geological Survey Professional Paper 422-I. US Geological Survey. US Government Printing Office. Washington.

Baker, E. A., K. E. Wehrly, P. W. Seelbach, L. Wang, M. J. Wiley, and T. P. Simon. 2005. A multimetric assessment of stream condition in the Northern Lakes and Forests Ecoregion using spatially explicit statistical modeling and regional normalization. Transactions of the American Fisheries Society. 134: 697-710.

Baker, M. E. and M. J. Wiley. 2009. Multiscale control of flooding and riparian-forest composition in Lower Michigan, USA. Ecology. 90: 145-159.

Baker, M. E., M. J. Wiley, and P. W. Seelbach. 2003. GIS-based models of potential groundwater loading in glaciated landscapes: considerations and development in Lower Michigan. Michigan Department of Natural Resources, Fisheries Research Report 2064. Ann Arbor.

Brenden, T. O., R. D. Clark, Jr., A. R. Cooper, P. W. Seelbach, L. Wang, S. S. Aichele, E. G.

Bissell, and J. S. Stewart. 2006. A GIS framework for collecting, managing, and analyzing multiscale landscape variables across large regions for river conservation and management. Pages 49–74 in R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Landscape influences on stream habitats and biological assemblages. American Fisheries Society Symposium 48. Bethesda, Maryland.

Bunn, S. E. and A. H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management. 30: 492-507.

Cho, W., Y. Park, H. Park, H. Kong, and T. Chon. 2011a. Ecological informatics approach to screening of integrity metrics based on benthic macroinvertebrates in streams. International Journal of Limnology. 47: S51-S62.

Cho, K. H., S. Sthiannopkao, Y. A. Pachepsky, K. Kim, and J. H. Kim. 2011b. Prediction of contamination potential of groundwater arsenic in Cambodia, Laos, and Thailand using artificial neural network. Water Research. 45: 5535-5544.

Diana, J. S. 2004. Biology and ecology of fishes, Second edition. Cooper Publishing Group. Traverse City, Michigan.

Draper, N. and H. Smith. 1981. Applied regression analysis. 2nd Edition. New York. John Wiley and Sons, Inc.

Dudgeon, D. 2000. Large-scale hydrological changes in tropical Asia: prospects for riverine biodiversity. BioScience. 50 (9): 793-806.

Dunne, T. and L. B. Leopold. 1978. Water in environmental planning. W.H. Freeman and Co., New York.

ESRI. 2005. ArcGIS. Version 9.1. (Build 722). Redlands, California.

Fausch, K. D., C. E. Torgersen, C. V. Baxter, and H. W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience. 52: 483-498.

Gordon, N. D., T. A. McMahon, B. L. Finlayson, C. J. Gippel, and R. J. Nathan. 2004. Stream hydrology: An introduction for ecologists. 2nd edition. John Wiley & Sons Ltd. ISBN 978-0-470-84385-1.

Hall, M. J., G. P. Closs, and R. H. Riley. 2001. Relationships between land use and stream invertebrate community structure in a South Island, coastal stream catchment. New Zealand journal of Marine and Freshwater Research. 35: 591-603.

Hamilton, D. A., R. C. Sorrell, and D. J. Holtschlag. 2008. A regression model for computing index flows describing the median flow for the lowest summer flow month in Michigan. U. S. Geological Survey, Scientific Investigations Report. 2008-5096. Reston, Virginia.

Hendrickson, G. E. and C. J. Doonan. 1972. Hydrology and recreation on the cold-water rivers of Michigan's southern peninsula. U.S. Geological Survey (in cooperation with the Michigan

Department of Natural Resources, Geological Survey Division. Water Information Series Report 3. Lansing, Michigan.

Higgins, J. V., M. Bryer, M. Lammert, and T. FitzHugh. 2005. A freshwater classification approach for biodiversity conservation planning. Conservation Biology. 19: 432-445.

Holtschlag, D. J. and H. M. Croskey. 1984. Statistical models for estimating flow characteristics of Michigan streams. U.S. Geological Survey. Water-Resources Investigations Report. 84-4207. Lansing. Michigan.

Hwang, S., N. Kim, S. Yoon, B. Kim, M. Park, K. You, H. Lee, H. Kim, Y. Kim, J. Lee, O. Lee, J. Shin, E. Lee, S. Jeon, and H. Joo. 2011. Distribution of benthic diatoms in Korean rivers and streams in relation to environmental variables. International Journal of Limnology. 47: S15-S33.

Hocking, R. R. 1976. The analysis and selection of variables in linear regression. Biometrics. 32.

Hyndman, R. J. and A. B. Koehler. 2006. Another look at measures of forecast accuracy. International Journal of Forecasting. 22 (4): 679-688.

Johnson R. K., M. T. Furse, D. Hering, and L. Sandin. 2007. Ecological relationships between stream communities and spatial scale: implications for designing catchment level monitoring programmes. Freshwater Biology. 52: 939-958.

Kilgour, B. W. and L. W. Stanfield. 2006. Hindcasting reference conditions in streams. Pages 623–639 in R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Influences of landscapes on stream habitats and biotic assemblages. American Fisheries Society. Symposium 48. Bethesda, Maryland.

KMA (Korea Meteorological Administration). Climate data of South Korea. 2011. Available: Http://www.kma.go.kr. (May 2011).

Lee, J., J. Han, H. K. Kumar, J. Choi, H. K. Byeon, J. Choi, J. Kim, M. Jang, H. Park, and K. An. 2011a. Relationship between three major stream assemblages and their environmental factors in multiple spatial scales. International Journal of Limnology. 47: S73-S89.

Lee, S., S. Hwang, J. Lee, D. Jung, Y. Park, and J. Kim. 2011b. Overview and application of the National Aquatic Ecological Monitoring Program (NAEMP) in Korea. International Journal of Limnology. 47: S3-S14.

Legendre, P. and P. Legendre. 1998. Numerical ecology. 2nd English edition. Elsevier Science. Amsterdam.

Leopold, L. B. 1968. Hydrology for urban land planning—a guidebook on the hydrologic effects of urban land use. U.S. Geological Survey. Circular 554. Washington, D. C.

Li, F., N. Chung, M. Bae, Y. Kwon, and Y. Park. 2012. Relationships between stream macroinvertebrates and environmental variables at multiple spatial scales. Freshwater Biology. 57 (10): 2107-2124.

Maeda-Veiga, A. and A. Sostoa. 2011. Observational evidence of the sensitivity of some fish species to environmental stressors in Mediterranean rivers. Ecological Indicators. 11: 311-317.

Meador, M. and D. M. Carlisle. 2007. Quantifying tolerance indicator values for common stream fish species of the United States. Ecological Indicators. 7: 329-338.

Montgomery, D. R. 1999. Process domains and the river continuum. Journal of the American Water Resources Association. 35: 397-410.

Moore, R. B., C. M. Johnston, K. W. Robinson, and J. R. Deacon. 2004. Estimation of total nitrogen and phosphorus in New England streams using spatially referenced regression models. U.S. Geological Survey. Scientific Investigations Report 2004-5012. Reston, Virginia.

Morley, S. A. and J. R. Karr. 2001. Assessing and restoring the health of urban streams in the Puget Sound Basin. Conservation Biology. 16: 1498-1509.

Naiman, R. J., S. E. Bunn, C. Nilsson, G. E. Petts, G. Pinay, and L. C. Thompson. 2002. Legitimizing fluvial ecosystems as users of water: an overview. Environmental Management. 30: 455-467.

Nash, J. E. and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models part I – A discussion of principles. Journal of Hydrology. 10 (3): 282-290.

NIER (National Institute of Environmental Research). 2009. Investigation and evaluation of aquatic ecosystem of South Korean streams and rivers. Final Report. May 2009. Ministry of Environment. S. Korea.

Park, K. S. Effects of different sampling effort and taxonomic resolution on assessment metrics. 2007. MS dissertation. University of Michigan. Ann Arbor, MI.

Petts, G. E., M. A. Bickerton, C. Crawford, D. N. Lerner, and D. Evans. 1999. Flow management to sustain groundwater-dominated stream ecosystems. Hydrological processes. 13: 497-513.

Piggott, A. R. and B. P. Neff. 2005. Calculation of streamflow statistics for Ontario and the Great Lakes states. U.S. Geological Survey. Open File Report 2005-1295. Reston, Virginia.

Poff, N. L. and J. V. Ward. 1989. Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Canadian Journal of Fisheries and Aquatic Sciences. 46: 1805-1818.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. C. Stromberg. 1997. The natural flow regime: a paradigm for river conservation and restoration. Bioscience. 47: 769-784.

Power, G., R. S. Brown, and J. G. Imhof. 1999. Groundwater and fish-insights from northern North America. Hydrological Processes. 13: 401-422.

Richter, B. D., R. Mathews, D. L. Harrison, and R. Wigington. 2003. Ecologically sustainable water management: managing river flows for ecological integrity. Ecological Applications. 13: 206-224.

Ries, K. G. III, P. A. Steeves, J. D. Coles, A. H. Rea, and D. W. Stewart. 2004. StreamStats: a U.S. Geological Survey web application for stream information. U.S. Geological Survey. Fact Sheet 2004-3115. Reston, Virginia.

Riseng, C., M. J. Wiley, and R. J. Stevenson. 2004. Hydrologic disturbance and nutrient effects on the structure of benthic stream communities: A covariance structure analysis. Journal of the North American Benthological Society. 23 (2): 309-326.

Riseng, C. M., M. J. Wiley, R. J. Stevenson, T. Zorn, and P. W. Seelbach. 2006. Comparison of Coarse Versus Fine Scale Sampling on Statistical Modeling of Landscape Effects and Assessment of Fish Assemblages of the Muskegon River, Michigan. In: Landscape influences on stream habitats and biological communities. American Fisheries Society Symposium 48. Bethesda, Maryland.

Riseng, C. R., M. J. Wiley, B. E. Pijanowski, P. W. Seelbach, and R. J. Stevenson. 2010. An ecological assessment of Great Lakes tributaries in the Michigan Peninsulas. Journal of Great Lakes Research. 36: 505-519.

Riseng, C. R., M. J. Wiley, R. W. Black, and M. D. Munn. 2011. Impacts of agricultural land use on biological integrity: a causal analysis. Ecological Applications. 21 (8): 3128-3146.

Seelbach, P. W., M. J. Wiley, M. E. Baker, and K. E. Wehrly. 2006. Initial classification of river valley segments across Michigan's Lower Peninsula. Pages 25-48 in R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Influences of landscapes on stream habitats and biotic assemblages. American Fisheries Society. Symposium 48. Bethesda, Maryland.

Seelbach, P. W., L. C. Hinz, M. J. Wiley, and A. R. Cooper. 2011. Use of multiple linear regression to estimate flow regimes for all rivers across Illinois, Michigan, and Wisconsin. Michigan Department of Natural Resources (DNR). Fisheries Division. Fisheries Research Report 2095. Ann Arbor, Michigan.

Simmons, D. L. and R. J. Reynolds. 1982. Effects of urbanization on base flow of selected south-shore streams, Long Island, New York. Water Resources Bulletin. 18: 797-805.

Smakhtin, V. U. 2001. Low flow hydrology: a review. Journal of hydrology 240: 147-186.

Sowa, S. P., G. Annis, M. E. Morey, and D. D. Diamond. 2007. A gap analysis and comprehensive conservation strategy for riverine ecosystems in Missouri. Ecological Monographs. 77: 301-334.

SPSS, Inc. 2003. SPSS for Windows Version 12.0. SPSS, Inc., Chicago, Illinois.

Stauffer, J. C., R. M. Goldstein, and R. M. Newman. 2000. Relationship of wooded riparian zones and runoff potential to fish community composition in agricultural streams. Canadian Journal of Fisheries and Aquatic Sciences. 57: 307-316.

Steen, P. J., D. R. Passino-Reader, and M. J. Wiley. 2006. Modeling Brook Trout Presence and Absence from Landscape Variables Using Four Different Analytical Methods. In: Landscape influences on stream habitats and biological communities. American Fisheries Society Symposium 48, 2006. Bethesda, Maryland.

Stevenson, R. J., S. T. Rier, C. M. Riseng, R. E. Schultz, and M. J. Wiley, 2006. Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria. Hydrobiologia 561: 149–165.

Thorp, J. H., M. C. Thoms, and M. D. Delong. 2005. The riverine ecosystem synthesis: biocomplexity in river networks across space and time. River Research and Applications. 22: 1-25.

Trush, W. J., S. M. McBain, and L. B. Leopold. 2000. Attributes of an alluvial river and their relation to water policy and management. Proceedings of the National Academy of Sciences. 97: 11858-11863.

U.S. ACE (U.S. Army Corps of Engineers). 2011. Hydrologic Engineering Center (HEC)-DSSVue 2. 0. 1. Available: Http://www.hec.usace.army.mil/software/hec-dssvue/.

Velleman, P. F. 1997. DataDesk Version 6.0: Statistics Guide. Data Description. Ithaka, NY.

WAMIS (Water Management Information System). 2011. The Ministry of Land, Infrastructure, and Transport. Stream flow gages and flow data. Republic of Korea. Available: Http://wamis.go.kr.

WAMIS (Water Management Information System). 2013. The Ministry of Land, Infrastructure, and Transport. Republic of Korea. Available: Http://wamis.go.kr.

Wang, L., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela, and P. Stewart. 2003. Watershed, Reach, and Riparian Influences on Stream Fish Assemblages in the Northern Lakes and Forest Ecoregion, USA. Canadian Journal of Fisheries and Aquatic Sciences. 60: 491-505.

Wang, L., P. W. Seelbach, and R. M. Hughes. 2006. Introduction to landscape influences on stream habitats and biological assemblages. Pages 1-24 in R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Influences of landscapes on stream habitats and biotic assemblages. American Fisheries Society. Symposium 48. Bethesda, Maryland.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 1997. Landscape-based models that predict July thermal characteristics of Lower Michigan Rivers. Michigan Department of Natural Resources (DNR). Fisheries Division. Fisheries Research Report 2037. Ann Arbor, MI.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 2006. Influence of landscape features on summer water temperatures in lower Michigan streams. Pages 113-128 In R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Influences of landscapes on stream habitats and biotic assemblages. American Fisheries Society. Symposium 48. Bethesda, Maryland.

Wiley, M. J., S. L. Kohler, and P. W. Seelbach. 1997. Reconciling landscape and site based views of stream communities. Freshwater Biology. 37: 133-148.

Wiley, M. J., P. W. Seelbach, K. E. Wehrly, and J. S. Martin. 2003. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. Pages 201-220 in T. P. Simon, editor. Biological response signatures: indicator patterns using aquatic communities. CRC Press. Boca Raton, Florida.

Wiley, M. J., B. C. Pijanowski, R. J. Stevenson , P. W. Seelbach, P. L. Richards, C. Riseng, D. W. Hynman, and J. M. Koches. 2008. Integrated modeling of the Muskegon River: Tools for ecological risk assessment in a Great Lakes Watershed. In: W. Ji, editor, Wetland and water resource modeling and assessment: A Watershed Perspective. Taylor & Francis, London.

Wiley M. J., D. W. Hyndman, B. C. Pijanowski, A. D. Kendall, C. Riseng, E. S. Rutherford, S. T. Cheng, M. L. Carlson, J. A. Tyler, R. J. Stevenson, P. J. Steen, P. L. Richards, P. W. Seelbach, J. M. Koches, and R. R. Rediske. 2010. A multi-modeling approach to evaluating climate and land use change impacts in a Great Lakes River Basin. Hydrobiologia. 657: 243-262.

Winter, T. C. 2001. The concept of hydrologic landscapes. Journal of the American Water Resources Association. 37: 335-349.

Winter, T. C., J. W. Harvey, O. L. Franke, and W. M. Alley. 1998. Ground water and surface water: a single resource. U.S. Geological Survey. Circular 1139. Reston, Virginia.

Yoon, J., J. Kim, M. Byeon, H. Yang, J. Park, J. Shim, H. Song, H. Yang, and M. Jang. 2011. Relationship between three major stream assemblages and their environmental factors in multiple spatial scales. International Journal of Limnology. 47: S91-S105.

Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2002. Distributions of stream fishes and their relationship to stream size and hydrology in Michigan's Lower Peninsula. Transactions of the American Fisheries Society. 131: 70-85.

Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2004. Utility of species-specific, multiple linear regression models for prediction of fish assemblages in rivers of Michigan's Lower Peninsula. Michigan Department of Natural Resources, Fisheries Research Report 2072, Ann Arbor.

Zorn, T. G., P. W. Seelbach, E. S. Rutherford, T. C. Wills, S. Cheng, and M. J. Wiley. 2008. A regional-scale habitat suitability model to assess the effects of flow reduction on fish assemblages in Michigan streams. Michigan Department of Natural Resources, Fisheries Research Report 2089, Ann Arbor.

Chapter 3 : The predictive performance evaluation of four different analytical methods in the estimation of Korean stream flow regimes from landscape variables

Abstract

The description and estimation of flow regimes in South Korean streams and rivers is an important step in accurately assessing aquatic ecosystems health and planning efficient water management strategies. In most countries, the network of active flow gauging stations is small relative to the scope of sampling undertaken in national-scale water quality monitoring programs. The main goal of this study was to evaluate four different analytical approaches to site-specific modeling of flow in gauged river segments. The analytical models compared included multiple linear regression (MLR), principal component regression (PCR), artificial neural networks (ANN), and the combination of principal components and an artificial neural networks (PC-ANN). I found that overall, each of the four methods did well at predicting Korean stream flows, although non-linear models showed slightly better accuracy than linear models across the range of high and low seasonal flows. Flows predicted by four different methods showed significantly high correlations ($p<0.01$) among them and with the observed flows in a validation dataset. This predictive performance comparison of analytical methods showed that the best choice may largely be based on convenience and familiarity with analytical methods, rather than predictive performance of each model in the prediction of Korean flow regimes.

Introduction

In 2008 the South Korean government initiated a Korean Nationwide Ecological Assessment Program (KNEAP) in order to maximize economic values of the country's freshwater streams and rivers (NIER 2009, Lee et al. 2011b, WAMIS 2011, Li et al. 2012). Despite a relatively large national gauging system, existing reports and publications from KNEAP exploring land use influences on aquatic ecosystems have generally ignored spatial variability in flow regimes as an important covariate (e.g., Bae et al. 2011, Cho et al. 2011a, Hwang et al. 2011, Lee et al. 2011a, Lee et al. 2011b, Yoon et al. 2011). For example, in KNEAP's most recent assessment Li et al. 2012 did not consider any stream related flow information in examination of the relationship between macroinvertebrates and environmental stressors. Since it has been well demonstrated that flow regimes play a critical role in shaping riverine communities (Poff et al. 1989, Poff et al. 1997, Stauffer et al. 2000, Bunn and Arthington 2002, Baker et al. 2003, Stevenson et al. 2006, Anonymous 2008, Zorn et al. 2008, Baker and Wiley 2009), failure to integrate hydrologic data into ecosystem assessments is certainly problematic, and stems largely from the lack of flow data for the vast majority of biological monitoring sites used by KNEAP.

In South Korea, as in most countries, hydrologic and biological river monitoring is carried out by completely different agencies and/or teams of researchers, so it is not surprising that data are collected usually at different sets of sites, and often with very different scales of spatial coverage. In order integrate hydrologic and biologic data sets in comprehensive analyses, hydrologic modeling of specific biological assessment sites is often the best alternative (Holtschlag and Croskey 1984, Smakhtin 2001, Hamilton et al. 2008, Seelbach et al. 2011).

Predicting site-specific flow regimes at un-gauged biological sampling sites is necessary to apply a more integrated analysis of the current ecological status of South Korean rivers and streams.

The site-specific estimation of stream flow regimes from landscape variables using multiple linear regression (MLR) analysis (e.g., Wiley et al 1997, Allan and Hinz 2004, Hamilton et al. 2008, Seelbach et al. 2011) follows naturally from the early work of hydrologists on linear hydraulic geometries (e.g., Dunne and Leopold 1978, Holtschlag and Croskey 1984). For example, Seelbach et al. 2011 predicted flow regimes for all rivers across Illinois, Michigan, and Wisconsin and examined the performance of MLR models at different spatial scales. Recently the State of Michigan (USA) has built site-specific enhanced regression estimation into its legal permitting process for pumped water withdrawals (Hamilton et al. 2008).

With advances in analytical technologies and tools, other modeling approaches have also become routinely available, and these are increasingly applied to the problem of hydrologic prediction. Empirical models (Obropta et al. 2007) are relatively easy to construct if adequate regional gauging data sets are available. Empirical approaches include the development of predictive functions taken from multiple linear regression, neural networks, and support vector machines including principal component regression (PCR). MLR and PCR approaches are not free from statistical assumptions, including normality, randomness, and the absence of outliers (Gros 1997, Cho et al. 2011b). In particular, inattentive contemplation of high correlations among independent variables can decrease the statistical robustness, eventually resulting in significant prediction errors (Mac Nally 2002, Cho et al. 2011b). Therefore, more advanced and nonlinear models are now being applied in many water resource problems, including hydrological process description, water quality modeling, and dam operation planning (Maier and Dandy 1996, Wen and Lee 1998, Lee et al. 2003, Riad et al. 2004, Sarangi and Bhattacharya 2005, Tayfur et al. 2005, Holmberg et al. 2006, Cho et al. 2011b).

Although these various analytical approaches have been used in environmental and ecological applications, direct performance comparisons of alternate analytical methods are rarely been made. Recently, water resource-related studies comparing alternate model performance have often observed that newer analytical technologies showed higher accuracy than simple linear models (Lek et al. 1996, Franklin 1998, Vayssieres et al. 2000, Steen et al. 2006, Cho et al. 2011b). However, a careful comparison of alternate empirical methods in the estimation of Korean stream flow regimes has not been previously conducted, and is undertaken here as a preliminary step towards a more integrated ecological assessment of South Korean streams and rivers.

My main objective was to evaluate the performance of four different modeling techniques (MLR, PCR, artificial neural networks (ANN), and the combination of principal component and artificial neural networks (PC-ANN)) in the estimation of Korean stream flow characteristics using landscape (GIS-extracted) variables. Our goal was to contribute to the discussion of the relative advantages and disadvantages of alternate methods for the estimation of stream flow regimes in general.

Materials and methods

Data collection and summary

Daily stream flow data from gauges across South Korean Peninsula were used in this study. Flow data from each gauge were obtained from the WAter Management Information System (WAMIS 2011), the National Institute of Environmental Research (NIER), the Ministry of Environment, Korea. Of the 603 listed discharge gauges, daily discharge data from 163 gauges (Figure 3.1A) were used eliminating 440 sites, based on the following four selection criteria. First, gauges that were non-operational in 2009 were eliminated because much of the biological data I analyzed in later chapters were obtained in 2009. Second, gauges with less than ten years of daily discharge data were removed since our interest was in modeling flow regimes, not specific dates. However, nineteen of these sites were intentionally included in order to balance gaging station distribution, in consideration of spatial stability of discharge data, even though they only had eight or nine years of discharge data. Third, gauges with discharges that were heavily impacted by anthropogenic activities were removed. Last, extreme outlier sites (11 out of 174 sites) were excluded based on boxplot and scatter plot assessments with landscape variables, using Datadesk 6.0 (Velleman 1997).

With qualified stream discharge records from the 163 gauges, I performed flow frequency analysis for each site using HEC-DSSVue 2.0.1 (U.S. ACE 2011), producing flow duration estimates and curves for the period of record. Summary statistics of three major percent exceedance discharges (10%, 50%, and 90%) were computed to represent high, median, and baseflow regime indicators, respectively (Table 3.1, Figure 3.2). The smaller percent exceedance discharges indicate high flow conditions for the data series, while the larger percent exceedance discharges indicate persistent or low flow conditions.

Candidate landscape attributes (independent variables) that were necessary to develop models were summarized at the catchment scales from digital maps of elevation, mean precipitation, mean air temperature, mean humidity, land cover/land use, and surficial geology (soil type and soil infiltration rate) using ArcGIS 9.1 (ESRI 2009). The digital maps of land cover/land use, surficial geology, and elevation were obtained from the WAter Management Information System (WAMIS 2011). The surficial geology maps included soil types and soil penetration rates. Soil penetration rates were categorized from very excellent with high penetration rate (category 1) to very poor with very low penetration rate (category 7). The digital contour maps of regional climate data (mean precipitation, mean air temperature, and mean humidity) were created with the observed data obtained from the Korea Meteorological Administration (KMA 2011), Korea (Figure 3.1B). The regional climate data included mean annual summary collected from 1981 to 2010 at 63 operational weather observation stations. Independent variables used in the predictive models were based on previous research (Wiley et al. 1997, Allan and Hinz 2004, Steen et al. 2006, Seelbach et al. 2011) and included catchment size, latitude, longitude, catchment slope, channel slope, mean precipitation, mean air temperature, mean humidity, number of dams, proportions of catchment land use, proportions of soil type, and proportions of soil infiltration rate. All independent variables in the models were transformed to natural log form after adding the integer 1 or 0.01 to the variable in order to maximize linearity within the modeled relationships and to meet assumptions of normality for all variables (Wiley et al. 2003, Riseng et al. 2006). Specifically, 1 was added to catchment area, precipitation, and site elevation and 0.01 was added to catchment slope, site slope, land cover/land use, and surficial geology. However, number of dams was not transformed in MLR models because the variable showed good linearity and normality without transformation.

Modeling approach: MLR, PCR, ANN, and PC-ANN

The decision to include independent variables in the MLR and PCR models to predict percent exceedance discharges was made with a manual, stepwise regression approach (Zorn et al. 2004, Allan and Hinz 2004, Steen et al. 2006, Seelbach et al. 2011) and the hierarchy and spatial scales of environmental factors (Allan and Castillo 2007) using Datadesk 6.0 (Velleman 1997) and SPSS 18.0 (SPSS Inc. 2009). Independent variables were inserted in the model in the following order: 1) catchment area, 2) precipitation, 3) catchment slope, 4) site elevation, 5) land cover/land use, 6) surficial geology variables (soil type or penetration rate), 7) land cover/land use variables, and 8) number of dams. The independent variables in the MLR models that would maximize R^2, be significant at $p < 0.05$, and have a t-ratio greater than 2 in the model were selected. However, a few important variables, significant at $p < 0.10$, were intentionally included, when they were considered to be causally critical for prediction of exceedance discharges. If independent variables in the model were highly correlated, the variable that best improved model fit was selected. The independent variables selected for the MLR and PCR models were then used in the ANN and PC-ANN models in order to produce a standard set of model structures that could be reasonably compared. MLR, PCR, ANN, and PC-ANN analyses were conducted using Datadesk 6.0 (Velleman 1997) and SPSS 18.0 (SPSS Inc. 2009).

The stream discharge gauging dataset was randomly partitioned into model-generation and model-test groups. A total of 30 gauges, approximately 18.4% of the total, were set aside as a model-test group and were used to evaluate and validate models. The remaining 133 sites (approximately 81.6%) were set into a model-generation group, which was used for building models for each exceedance discharge.

Multiple linear regressions (MLR)

Each MLR model can be defined as follows,

$$\ln(y) = \beta_0 + \sum_{i=1}^{n} \beta_i \ln(x_i) \quad (1)$$

where x_i is an explanatory variable i (i.e., catchment size, climate properties, land uses, soil properties, and site-specific stream properties), y is the response variable (Q10, Q50, and Q90), β_i is the regression coefficient of explanatory variable x_i, and β_0 is the value of the intercept.

Principal component regressions (PCR)

Principal component regression is a regression analysis that uses principal component analysis (PCA) when calculating regression coefficients (Indahl and Naes 1998, Park 1981, Bair et al. 2006). PCR approach has two advantages. First, multi-collinearity can be avoided by using PCs in place of the original variables. Second, the dimensionality of the regression is minimized by using only a subset of PCs. In the process of the PCR, an orthogonal linear transform of the original data generates a new set of variables (the principal components, PCs) and PC scores. The generated PCs and PC scores are then used in the regression as explanatory variables to estimate response variables.

Artificial neural network (ANN)

ANN is an advanced computational model that uses the interconnection of software neurons that can estimate values from inputs by feeding information through the network (Lawrence 1994, Bishop 1995, Cho et al. 2011b). The neurons in the different layers of each system are interconnected in the ANN model. An ANN model includes an input layer, hidden layer, and output layer. The input layer nodes take the input vectors and transfer the signals to the next layer. This process will be continued until the signals reach to the output layer and more detailed computational processes are well described in Norgaard et al. 2000 and Cho et al. 2011b.

Principal component-artificial neural network (PC-ANN)

PC-ANN merges PCA decomposition with ANN (Sousa et al. 2007). The PC-ANN takes the benefits of both analytical modeling approaches. The main difference between this approach and ANN is that PC scores generated from the orthogonal linear transformation of the original data are used as the input variables of ANN; other procedures for the optimization, training, and validation are the same as those for the ANN model.

Performance evaluation of four different predictive models

In order to evaluate the performance of four different predictive models, the mean absolute error (MAE; Hyndman and Koehler 2006), Nash-Sutcliffe model efficiency (NSE; Nash and Sutcliffe 1970) coefficient, variance inflation factor (VIF; Longnecker and Ott 2004), and Pearson product-moment correlation coefficient (Pearson’s r; Pearson 1895) were estimated with the observed and predicted flow discharges of three percent exceedance flows.

The Mean absolute error (MAE) is a statistical approach used to measure how close predicted values are to the observed values and can be defined as follows,

$$\text{MAE}=\left[n^{-1}\sum_{i=1}^{n}\left| x_{obs}-x_{pre}\right| \right] \tag{2}$$

where n indicates the number of observations of stream discharge for each percent exceedance. Here, x_{obs} and x_{pre} indicate the observed and predicted stream discharges, respectively.

The Nash-Sutcliffe model efficiency coefficient (NSE) was also computed to evaluate the predictive power of four different models using the predicted and observed stream discharges. If NSE value is greater than 0.5, the model showed acceptable accuracy. If NSE value is greater than 0.7, the model is in a good agreement with observation (Moriasi et al. 2007). The NSE can be defined as follows,

$$\text{NSE}=1-\frac{\sum_{t=1}^{T}(x_{obs}^{t}-x_{pre}^{t})^{2}}{\sum_{t=1}^{T}(x_{obs}^{t}-\overline{x}_{obs}^{t})^{2}} \tag{3}$$

where x_{obs} is the observed stream discharges, x_{pre} is the predicted stream discharges, and $\overline{x}_{obs}^{t}$ is the averaged value of the observed stream discharges.

The variance inflation factor (VIF) is the measurement of the severity of multi-collinearity in a regression analysis and is used for the examined linear models (MLR and PCR) to identify any collinearity problem in a matrix. If VIF is greater than 5, then it can be assumed that multicollinearity is high. Also, the proposed cut off value of VIF is 10 (Bowerman and O'Connell 1990, o’Brien 2007). Collinearity can sometimes results in statistical stability problems, such as high variance in estimated coefficients in the regression model. Therefore, it is necessary to investigate the VIF value in order to prevent colinearity in the MLR. However, this procedure is not necessary in PCR because VIF is always supposed to be 1, implying that all PCs are orthogonal; VIF refers to the effect of collinearity on the variance of the estimated coefficients (Longnecker and Ott 2004), such that

$$\mathrm{VIF}_i = \frac{1}{1 - R_i^2} \tag{4}$$

where R_i^2 is the multiple correlation coefficient between the ith explanatory variable and other explanatory variables in the regression model, and VIF_i is the variance inflation factor associated with the ith explanatory variable.

Results

All four analytical models were successfully developed and used to predict percent exceedance discharges. The overall predictive performances were clearly acceptable (Table 3.2) in all cases. The performance evaluation statistics demonstrated that both linear and non-linear models produced statistically satisfactory results, showing relatively high NSE values ranging from 0.77 to 0.88 in the model generation step and ranging from 0.70 to 0.85 in the test step. The average MAEs and NSEs for the overall models were 46.61 and 0.82 in the model test step and 48.57 and 0.76 in the test step, respectively.

Two-tailed Pearson correlation test also showed significantly high correlation among observed and all predicted discharges (Table 3.3). Overall Pearson correlation values ranged from 0.880 to 1.000 with mean of 0.945. The mean of Pearson r (0.975) among predicted discharges was relatively higher than the mean r (0.945) between observed and predicted discharges, indicating the reflection of of landscape variable influences in models.

Summary statistics of response and explanatory variables

The percent exceedance discharges (Q10, Q50, and Q90) represent high, median, and low stream flow events, respectively (Table 3.1 and Figure 3.2). One way ANOVA with post-hoc Tukey tests indicated that across the gauging dataset, 10% exceedance discharges (mean = 207.48cms) were significantly different from 50% (mean = 55.53cms) and 90% (mean = 19.21cms) exceedance discharges ($p<0.01$). However, 50% exceedance discharges were not significantly different from 90% exceedance discharges ($p>0.05$) reflecting the large variability of flow and water yields in S. Korea.

Landscape (explanatory) variables ultimately retained for the various models included drainage area, mean annual precipitation, urban land use, number of dams, and soil penetration rates (Table 3.1).Variables included in models for all three percent exceedance discharges were drainage area and mean annual precipitation. The drainage areas of stream flow gauges in South Korean streams and rivers varied from 50 km^2 to 23,316 km^2 with the mean of 3,080 km^2 and the mean annual precipitation varied from 1,072 mm to 1,588 mm with the mean of 1,319 mm. Urban land use varied from 0.00% to 48.80% with the mean of 7.44%. Number of dams was also an important variable for 50% and 90% exceedance discharges and ranged from 0 to 11 (mean = 1.48).

Predictive models for 10% exceedance discharge (high flows)

The predictive performance of the four different models for the 10% exceedance (high flow) discharge was compared using mean absolute percentage errors and Nash-Sutcliffe model efficiency coefficients (Table 3.2 and Figure 3.3). Overall, the NSE values indicated that the four different models all showed good predictive performance; ranging from 0.78 to 0.82 for the generation step and ranging from 0.72 to 0.77 for the test step. In general, the non-linear models (ANN and PC-ANN) had slightly lower error rates than the linear models (MLR and PCR) in the both generation and test steps. The ANN model had the highest NSE value (0.82) and the lowest MAE value (46.43) in the generation step among the four analytical models.

The landscape variables included in predictive models for the 10% exceedance discharges were drainage area, mean annual precipitation, and soil penetration rate category 1 (Table 3.4). The standardized regression coefficients of the MLR model for the 10% exceedance discharges indicated that drainage area was the most influential variable and followed by mean annual precipitation and soil penetration rate category 1, respectively. The average variance inflation factor (VIF) values for the MLR model ranged from 1.004 to 1.141 with the average of 1.095, suggesting a small problem with collinearity (Table 3.5).

Predictive models for 50% exceedance discharge (median flows)

Using the NSE criteria, median stream flows were also well predicted by all four analytical methods (Table 3.2 and Figure 3.4). The 50% models performed better across the board than 10% exceedance discharge models with the NSE ranges from 0.81 to 0.88 in the generation step and ranges from 0.70 to 0.75 in the test step. Also, MAE values of the 50% exceedance discharge models were much lower than those of the 10% exceedance discharge models in the generation step, although the MAE values of the test step showed opposite results. Similarly to the high flow modeling, the ANN and PC-ANN models for median flow had higher prediction accuracy than linear models in the generation data set analysis, although the situation was reversed in my analysis of the validation data set. No significant difference was observed between the performance of the MLR and PCR models.

The predictive models for the 50% exceedance discharge included drainage area, mean annual precipitation, urban land use, soil penetration rate, and number of dams (Table 3.4). The standardized regression coefficients of the MLR model indicated that drainage area was the most influential variable and followed by number of dams, urban land use, mean annual precipitation, and soil penetration rate category 6, respectively. Collinearity issues in the MLR increased relative to the high flow model, but the values still suggest only a small problem with collinearity (Table 3.5).

Predictive models for 90% exceedance discharge (low flows)

Each method also reasonably predicted 90% exceedance discharge with only minor differences in NSE statistics. The NSE values for the ANN and PC-ANN, the non-linear models, were 0.85 and 0.82 in the generation step and 0.83 and 0.84 in the test step, respectively. The NSE values for MLR and PCR were 0.81 and 0.77 in the generation step and 0.84 and 0.85 in the test step (Table 3.2 and Figure 3.5). However, in terms of error rate the non-linear models again out-performed the linear models in the generation set analyses.

The 90% exceedance discharge had the same independent variables in the predictive models and included drainage area, mean annual precipitation, urban land use, soil penetration rate, and number of dams (Table 3.4). However, the standardized regression coefficients of the MLR model indicated that the importance of independent variables in the MLR model was slightly different from the 50% exceedance discharge model. Drainage area was the most influential variable and followed by number of dams, mean annual precipitation, urban land use, and soil penetration rate category 6, respectively.

Discussion

As quantitative hydrologic information has been crucially important for the effective conservation and management of freshwater resources and ecosystems (Trush et al. 2000, Fausch et al. 2002, Zorn et al. 2002, Diana 2004, Allan and Castillo 2007), flow regime analysis and modeling is a common practice in aquatic ecosystem conservation and research (Petts et al. 1999, Zorn et al. 2002, Allan and Hinz 2004, Piggott and Neff 2005, Hamilton et al. 2008, Seelbach et al. 2011). South Korean aquatic ecosystems suffer from extreme high flow events in each monsoon season, while flows in the rest of seasons are typically quite low. The South Korean peninsula has a dramatic monsoon climate with heterogeneous geomorphology. Mountainous areas create higher stream slopes and relatively short stream channel lengths. Conversely coastal lowland channels are typically mild in slope and longer in length. The combination of strong seasonality in climate and landscape heterogeneity makes prediction of stream flows from empirical data a challenge. And while all of the methods explored produced reasonably good models, accuracy here was somewhat less than has been reported by workers in some other geographic settings. For example, while my MLR models had R^2 values ranging from 0.78 to 0.82; Seelbach et al (2011) working in the Great Lakes basin of North America reported values of 0.96-0.98 for Michigan MLR models, 0.74-0.98 for Illinois, and 0.94-0.98 for Wisconsin.

All high, median, and low flow models consistently included drainage area and mean annual precipitation as the most important predictive variables (Table 3.4). Urban land-use, soil properties, and number of dams were also often important variables. High stream flows (Q90) were mainly predicted by drainage area and mean annual precipitation; whereas the importance of number of dams and urban land-use was relatively stronger for low exceedance flows

(Q90)than higher exceedances. These patterns are consistent with reports in previous studies by Hamilton et al. 2008 and Seelbach et al. 2011.

Linear regression models (MLR and PCR)

Linear regression modeling of stream flow has a long history and is a well established method in both hydrologic and ecological studies; the various advantages and disadvantages of MLR for this purpose have been well described elsewhere (Holtschlag and Croskey 1984, Allan and Hinz 2004, Hamilton et al. 2008, Seelbach et al. 2011). In our study, both MLR and PCR methods had satisfactory MAE and NSE values (Table 3.2). The performance of MLR was slightly better than PCR in the Q50 and Q90 models, although for the Q10 there was almost no difference in MAE and NSE values. There is no doubt that linear regression models can be used for accurate stream flow estimation.

Linear regression models (MLR and PCR) generally have better performance with continuous response variables than dichotomous or categorical variables (Zar 1999, Steen et al. 2006). However, MLR can suffer from issues of multi-collinearity. Variance inflation factor (VIF) is a useful index to check for multi-collinearity among explanatory variables (Longnecker and Ott, 2004). Minor multi-collinearity was observed in MLR models, but the VIF values were usually within acceptable ranges (Table 3.5). However, I found no multi-collinearity problem in PCR models although predictive performance was slightly lower than MLR models. Thus, the appropriate selection between two linear analytical methods can be carefully considered with multi-collinearity of explanatory variables.

Non-linear models (ANN and PC-ANN)

Both non-linear models (ANN and PC-ANN) showed good predictive performance with relatively lower MAEs and higher NSEs than those of linear regression models (Table 3.2). Between non-linear models, ANN models showed slightly better predictive performance than PC-ANN models. Similar results were reported in a study of the prediction methods for groundwater arsenic (Cho et al. 2011b). However, this was not the case in comparisons of modeling approaches for brook trout presence/absence in Michigan rivers from landscape variables (Steen et al. 2006). It seems likely that the selection of best predictive model will depend on the specific on data types and distributions of response and explanatory variables involved.

Neural networks are distribution-free and working well with messy data and nonlinear responses (Bishop 1995, Gurney 1997, Norgaard 2000, Holmberg et al. 2006, Steen et al 2006). The ANN approach will be relatively effective and easier for researchers with less ecological and statistical background, when they have difficulty in choosing explanatory variables for the best predictive model. Previous studies showed that full neural network models with all variables and pruned neural network models produced satisfactory results in predicting the generation and test sets (Steen et al. 2006). Despite of these advantages, neural networks are fairly new methods and not familiar to many researchers. ANN consists of the interconnections between neurons (processing elements or units) in the different layers of each system. It is not simple for beginners to understand the process and interpretation of process layers. Thus, it may not be practical to interpret the relationship between response and explanatory variables.

Error in databases and other limitations

The selection of explanatory variables for four different analytical models was based on a manual, stepwise regression approach (Zorn et al. 2004, Allan and Hinz 2004, Steen et al. 2006, Seelbach et al. 2011) and the hierarchy and spatial scales of environmental factors (Allan and Castillo, 2007) in MLR models. Once this was completed, the selected explanatory variables were applied to the other three predictive models in order to build standard forms and examine the comparative performance of four different predictive models. Although all predictive models all showed satisfactory prediction results, they did not necessarily represent the "best-fit" model with the optimal explanatory variables for each predictive method, and statistical accuracy may differ with different variable selections (Cho et al. 2009). Thus, it might be worthwhile to compare the predictive performance of the "best-fit" model using each method. However, comparing different methods employing different independent variables would make comparisons difficult to evaluate. In this study I focused on the performance of standard forms of different predictive models to help clarify differences due to methodology alone.

The response and explanatory data used for this comparison study were collected at different time scales, which were available at the time of data collection. These different time periods might cause some error in the predictive models. Exceedance discharges for each site included at least ten years of data, but many sites included more than twenty or thirty years of stream flow data. The data for mean annual precipitation, mean annual air temperature, and mean annual humidity were collected from 1981 to 2010. Land-use, soil properties, and number of dams were based on GIS data measured in 2009. However, I believe that these data can generally represent long-term average trends in Korean stream flows, as other similar studies also successfully employed this long term data summary (Wehrly et al. 1997, Wiley et al. 2003, Riseng et al. 2010, Wiley et al. 2010, Riseng et al. 2011, Seebach et al. 2011). Also, the use of long term data summaries can avoid errors that might be induced by short term year-to year variation in either stream flow or climate data.

Overall, all four predictive models successfully estimated the Korean stream flows for three (high, median, and low) representative annual stream flow events. When predictive performance was evaluated with several statistical indices, the order for accuracy was as follows: ANN, PC-ANN, MLR, and PCR. Non-linear models showed slightly better prediction than linear models. However, model predictions showed only small differences relative to each other, and all predicted flows were highly correlated. Thus, the selection of an appropriate model might be based on other criteria than this predictive accuracy.

Table 3.1. Summary statistics of dependent and independent variables used for comparison of predictive models. All other variables not used for model construction were removed from the table. SD indicates standard deviation.

Variables	Short code	n	Mean	Median	SD	Min	Max
Dependent variables							
10 % exceedance discharge (cms)	Q10	163	207.48	44.90	461.87	1.69	3820.90
50 % exceedance discharge (cms)	Q50	163	55.53	6.55	176.71	0.18	1857.80
90 % exceedance discharge (cms)	Q90	163	19.21	1.58	52.26	0.03	323.90
Independent variables							
Drainage area (km^2)	DAREA	163	3080.42	487.54	5654.57	50.23	23316.70
Mean annual precipitation (mm)	MAPPT	163	1319.17	1323.47	98.09	1072.78	1588.31
Soil penetration rate 1 (proportion)	SOILP1	163	0.47412	0.48763	0.18100	0.00000	1.00000
Soil penetration rate 6 (proportion)	SOILP6	163	0.00061	0.00000	0.00313	0.00000	0.03564
Urban land use (proportion)	URBAN	163	0.07439	0.04000	0.09512	0.00000	0.48800
Number of dams	NDAMS	163	1.48	0.00	2.67	0.00	11.00

Table 3.2. Mean absolute percentage errors (MAEs) and Nash-Sutcliffe model efficiency coefficient (NSE) of each model for three percent exceedance discharges (Q10, Q50, and Q90).

		Generation steps		Test step	
		MAE	NSE	MAE	NSE
Q10	MLR	51.94	0.78	49.72	0.73
	PCR	51.94	0.78	49.69	0.73
	ANN	46.43	0.82	48.09	0.72
	PC-ANN	50.05	0.80	44.89	0.74
Q50	MLR	47.33	0.82	56.09	0.75
	PCR	49.07	0.81	57.11	0.72
	ANN	38.32	0.88	61.33	0.71
	PC-ANN	41.31	0.87	58.22	0.70
Q90	MLR	46.53	0.81	40.32	0.84
	PCR	53.02	0.77	35.72	0.85
	ANN	39.00	0.85	43.95	0.83
	PC-ANN	44.40	0.82	37.68	0.84

Table 3.3. Two tailed Pearson correlation tests among observed and predicted stream discharges for all sites combined (n=163). Bold indicates significance at p≤0.05, and bold and italics indicate significance at p≤0.01.

Q10					
	Observed	Predicted (MLR)	Predicted (PCR)	Predicted (ANN)	Predicted (PCANN)
Observed	1.000				
Predicted (MLR)	***0.880***	1.000			
Predicted (PCR)	***0.880***	***1.000***	1.000		
Predicted (ANN)	***0.899***	***0.972***	***0.972***	1.000	
Predicted (PCANN)	***0.891***	***0.981***	***0.981***	***0.982***	1.000
Q50					
	Observed	Predicted (MLR)	Predicted (PCR)	Predicted (ANN)	Predicted (PCANN)
Observed	1.000				
Predicted (MLR)	***0.901***	1.000			
Predicted (PCR)	***0.895***	***0.996***	1.000		
Predicted (ANN)	***0.925***	***0.960***	***0.954***	1.000	
Predicted (PCANN)	***0.919***	***0.970***	***0.967***	***0.984***	1.000
Q90					
	Observed	Predicted (MLR)	Predicted (PCR)	Predicted (ANN)	Predicted (PCANN)
Observed	1.000				
Predicted (MLR)	***0.902***	1.000			
Predicted (PCR)	***0.884***	***0.978***	1.000		
Predicted (ANN)	***0.919***	***0.970***	***0.951***	1.000	
Predicted (PCANN)	***0.906***	***0.984***	***0.967***	***0.982***	1.000

Table 3.4. Regression coefficients and collinearity statistics of the MLR and PCR models for 10%, 50%, and 90% exceedance discharges (Q10, Q50, and Q90). Bold unstandardized coefficients indicate significance at p ≤ 0.05, and bold and italics unstandardized coefficients indicate significance at p ≤ 0.01.

Independent variable (*i*)	Percent exceedance discharge								
	Q10			Q50			Q90		
	Unstandardized Coefficients		Standardized Coefficients	Regression Coefficients		Standardized Coefficients	Regression Coefficients		Standardized Coefficients
	b_i	Std. Error (SE_{bi})	Beta	b_i	Std. Error (SE_{bi})	Beta	b_i	Std. Error (SE_{bi})	Beta
R^2: 0.779				**R^2: 0.821**			**R^2: 0.807**		
Constant	**-16.368**	6.476		***-18.434***	6.183		***-19.442***	5.988	
ln(DAREA)	***0.865***	0.042	0.912	***0.780***	0.064	0.794	***0.560***	0.062	0.610
ln(MAPPT)	**2.006**	0.887	0.100	***2.712***	0.841	0.130	***2.989***	0.815	0.154
ln(SOILP1)	**-0.284**	0.119	-0.099						
ln(SOILP6)				**0.703**	0.343	0.078	***0.872***	0.332	0.104
ln(URBAN)				***0.252***	0.072	0.132	***0.194***	0.070	0.109
NDAMS				**0.093**	0.040	0.150	***0.200***	0.039	0.347
R^2: 0.779				**R^2: 0.814**			**R^2: 0.770**		
Constant	***4.049***	0.063		***2.468***	0.061		***1.529***	0.063	
PC1	***1.015***	0.063	0.663	***1.222***	0.061	0.770	***1.113***	0.063	0.752
PC2	-0.039	0.063	-0.026	**0.148**	0.061	0.093	**0.148**	0.063	0.100
PC3	***0.890***	0.063	0.582	***-0.189***	0.061	-0.119	**-0.129**	0.063	-0.087
PC4				***0.553***	0.061	0.349	***0.532***	0.063	0.359
PC5				***-0.438***	0.061	-0.276	***-0.358***	0.063	-0.242

Table 3.5. Collinearity statistics for MLR and PCR models. Variance inflation factors (VIFs) and average VIFs were summarized in order to see multicollinearity issues among independent variables used in MLR and PCR models.

Independent variable (i)	Percent exceedance discharge					
	Q10		Q50		Q90	
	VIF	Average VIF	VIF	Average VIF	VIF	Average VIF
DAREA	1.141		3.008		3.008	
MAPPT	1.140		1.160		1.160	
SOILP1	1.004	1.095		1.847		1.847
SOILP6			1.033		1.033	
URBAN			1.024		1.024	
NDAMS			3.008		3.008	
PC1	1.000		1.000		1.000	
PC2	1.000		1.000		1.000	
PC3	1.000	1.000	1.000	1.000	1.000	1.000
PC4			1.000		1.000	
PC5			1.000		1.000	

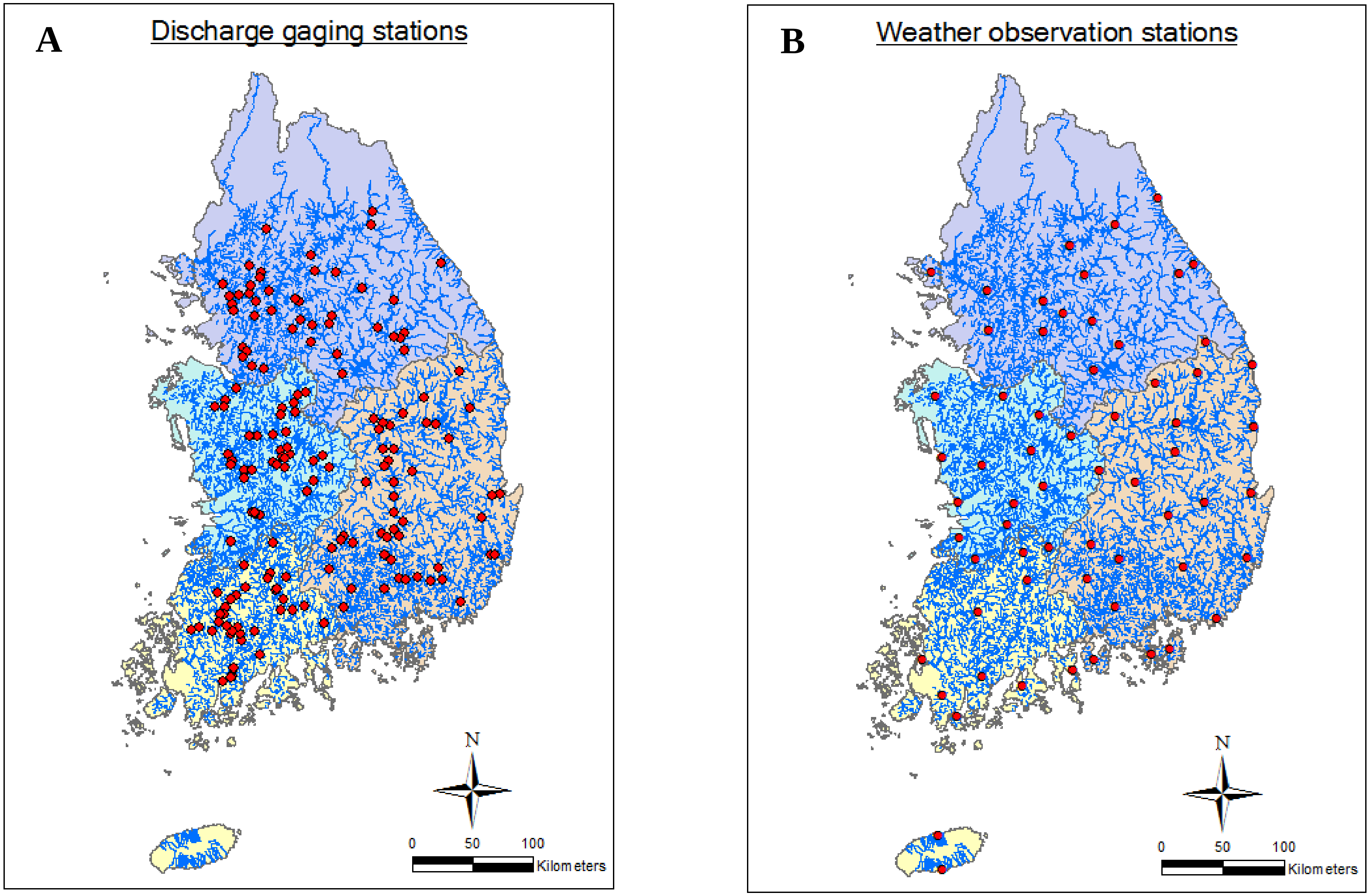


Figure 3.1. Locations of A) stream discharge gages (163 sites) and B) weather-observation stations (63 sites).

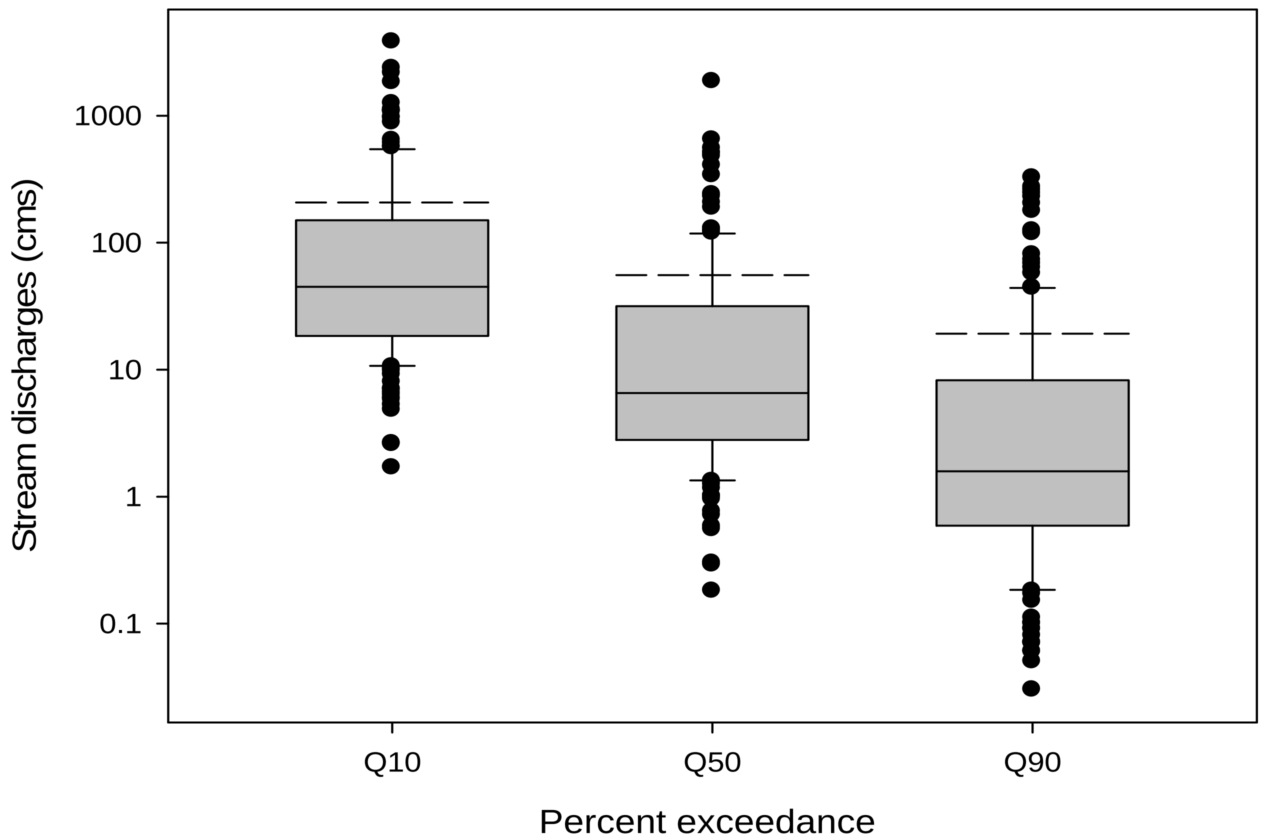


Figure 3.2. Boxplots showing the mean (dashed line), median (solid line), and range of stream discharges (cms) for exceedance discharges (Q10, Q50, and Q90) used for predictive models. Y-axis is log scaled for better view of discharge distribution.

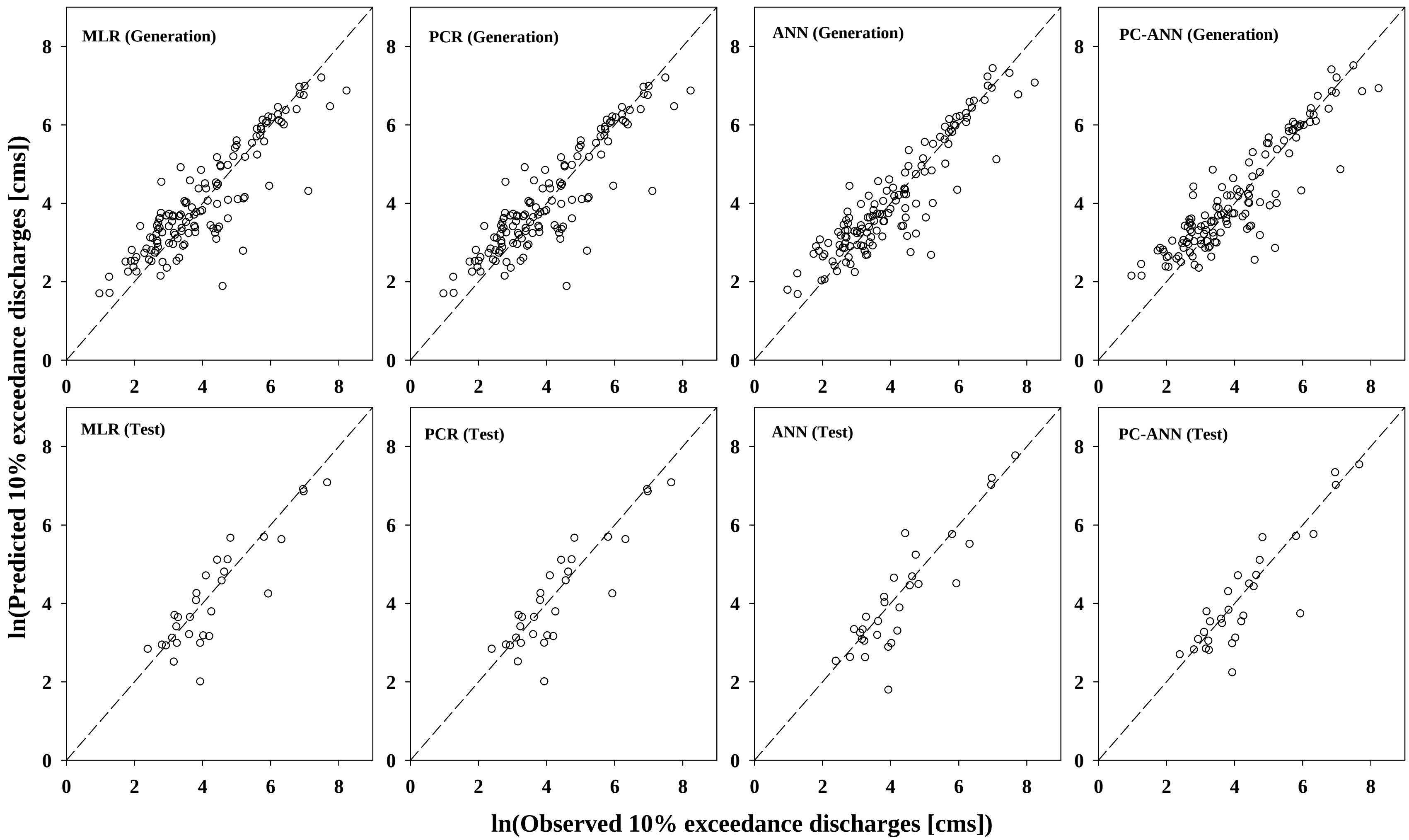


Figure 3.3. Scatter plots of observed and predicted 10% exceedance flows (Q10). Predicted exceedance flows were modeled by ANN, MLR, PC-ANN, and PCR.

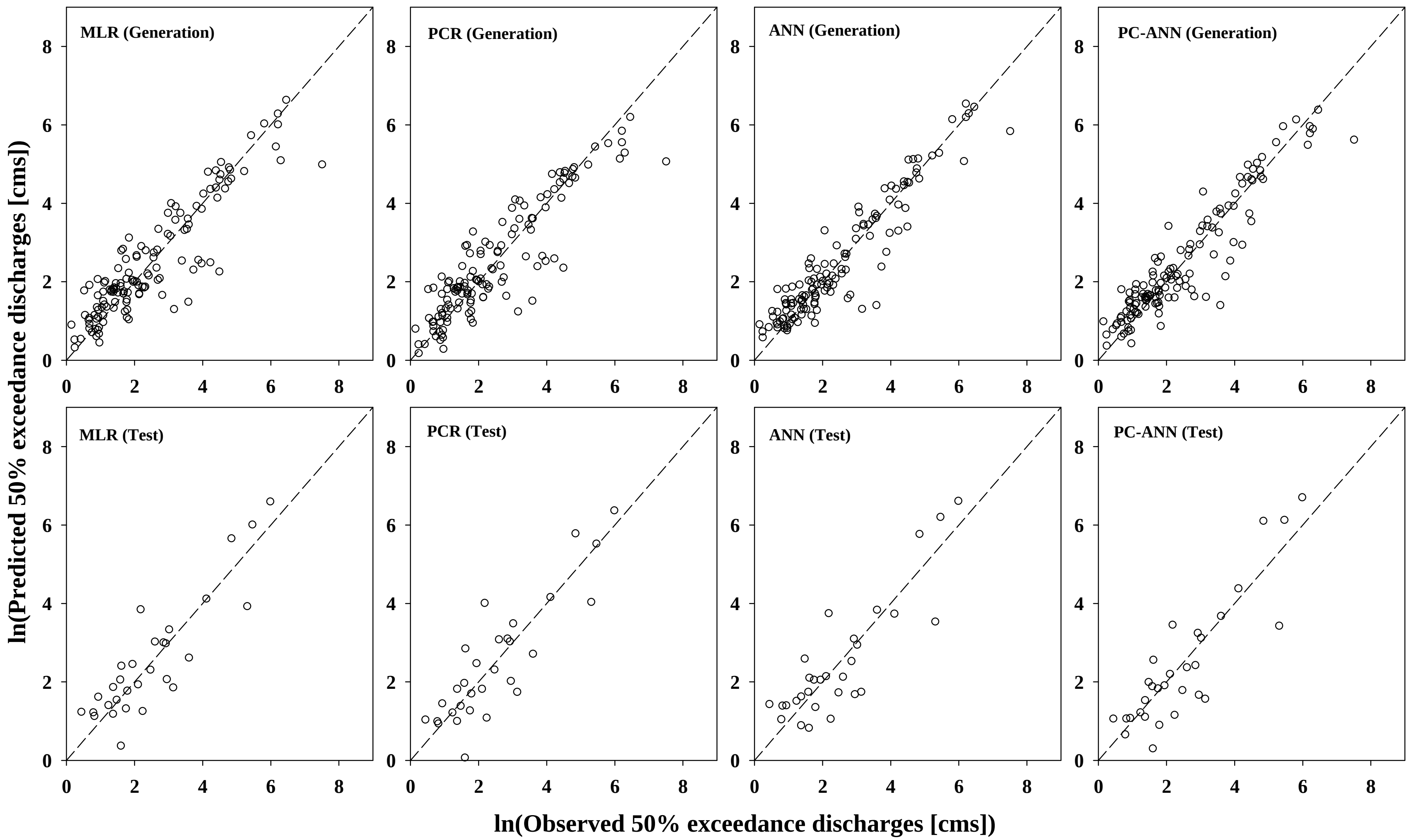


Figure 3.4. Scatter plots of observed and predicted 50% exceedance flows (Q50). Predicted exceedance flows were modeled by ANN, MLR, PC-ANN, and PCR.

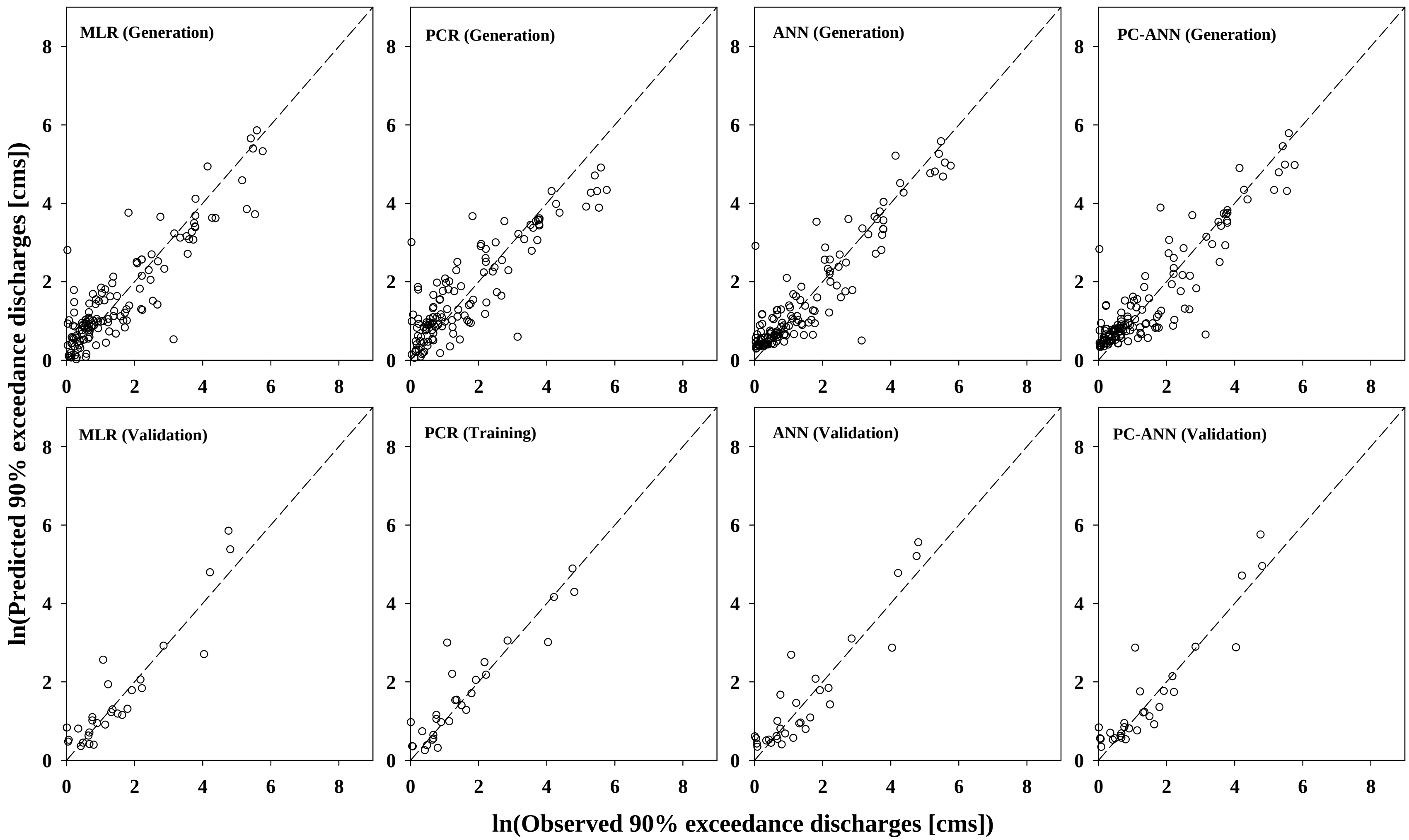


Figure 3.5. Scatter plots of observed and predicted 90% exceedance flows (Q90) Predicted exceedance flows were modeled by ANN, MLR, PC-ANN, and PCR.

Literature cited

Allan, D. and L. Hinz. 2004. An assessment of flows for rivers of the Great Lakes Basin. Final report to the Great Lakes Protection Fund, Chicago, Illinois.

Allan, J. D. and M. M. Castillo. 2007. Stream ecology: structure and function of running waters. Second edition. Springer. Dordrecht, The Netherlands.

Anonymous. 2008. Ecological limits of hydrologic alteration: environmental flows for regional water management. ELOHA Working Group. Fact Sheet. Available: http://www.nature.org/ELOHA. (May 2009).

Bae, M., Y. Kwon, S. Hwang, T. Chon, H. Yang, I. Kwak, J. Park, S. Ham, and Y. Park. 2011. Relationship between three major stream assemblages and their environmental factors in multiple spatial scales. International Journal of Limnology. 47: S91-S105.

Baker, M. E. and M. J. Wiley. 2009. Multiscale control of flooding and riparian-forest composition in Lower Michigan, USA. Ecology. 90: 145-159.

Baker, M. E., M. J. Wiley, and P. W. Seelbach. 2003. GIS-based models of potential groundwater loading in glaciated landscapes: considerations and development in Lower Michigan. Michigan Department of Natural Resources, Fisheries Research Report 2064. Ann Arbor.

Bair, E., T. Hastie, D. Paul, and R. Tibshirani (2006). Prediction by Supervised Principal Components. Journal of the American Statistical Association. 101 (473): 119-137.

Bishop, C. M. 1995. Neural Networks for Pattern Recognition. Oxford: Oxford University Press. ISBN 0-19-853849-9 (hardback).

Bowerman, B. L., and R. T. O'Connell. 1990. Linear statistical models: An applied approach. 2nd edition. Belmont, CA, Duxbury Press.

Bunn, S. E. and A. H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management. 30: 492-507.

Cho, K. H., J. Kang, S. J. Ki, Y. Park, S. M. Cha, and J. H. Kim. 2009. Determination of the optimal parameters in regression models for the prediction of chlorophyll-a: A case study of the Yeongsan Reservoir, Korea. Science of the Total Environmen. 407: 2536-2545.

Cho, W., Y. Park, H. Park, H. Kong, and T. Chon. 2011a. Ecological informatics approach to screening of integrity metrics based on benthic macroinvertebrates in streams. International Journal of Limnology. 47: S51-S62.

Cho, K. H., S. Sthiannopkao, Y. A. Pachepsky, K. Kim, and J. H. Kim. 2011b. Prediction of contamination potential of groundwater arsenic in Cambodia, Laos, and Thailand using artificial neural network. Water Research. 45: 5535-5544.

Diana, J. S. 2004. Biology and ecology of fishes, Second edition. Cooper Publishing Group. Traverse City, Michigan.

Dunne, T. and L. B. Leopold. 1978. Water in environmental planning. W.H. Freeman and Co., New York.

ESRI. 2009. ArcGIS. Version 9.1. (Build 722). Redlands, California.

Fausch, K. D., C. E. Torgersen, C. V. Baxter, and H. W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience. 52: 483-498.

Franklin, J. 1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. Journal of Vegetation Science. 9: 733-748.

Gros, P. 1997. Practical linear regression modelling. Basic principles. Validation methods. Oceanis. 23 (3): 359-515.

Gurney, K. 1997. An introduction to neural networks. London: Routledge. ISBN 1-85728-673-1 (hardback).

Hamilton, D. A., R. C. Sorrell, and D. J. Holtschlag. 2008. A regression model for computing index flows describing the median flow for the lowest summer flow month in Michigan. U. S. Geological Survey, Scientific Investigations Report. 2008-5096. Reston, Virginia.

Holtschlag, D. J. and H. M. Croskey. 1984. Statistical models for estimating flow characteristics of Michigan streams. U.S. Geological Survey. Water-Resources Investigations Report. 84-4207. Lansing. Michigan.

Hwang, S., N. Kim, S. Yoon, B. Kim, M. Park, K. You, H. Lee, H. Kim, Y. Kim, J. Lee, O. Lee, J. Shin, E. Lee, S. Jeon, and H. Joo. 2011. Distribution of benthic diatoms in Korean rivers and streams in relation to environmental variables. International Journal of Limnology. 47: S15-S33.

Holmberg, M., M. Forsius, M. Starr, M. Huttunen. 2006. An application of artificial neural networks to carbon, nitrogen and phosphorus concentrations in three boreal streams and impacts of climate change. Ecological Modelling. 195 (1-2): 51-60.

Hyndman, R. J. and A. B. Koehler. 2006. Another look at measures of forecast accuracy. International Journal of Forecasting. 22 (4): 679-688.

Indahl, Ulf G. and T. Naes. 1998. Evaluation of alternative spectral feature extraction methods of textural images for multivariate modeling. Journal of Chemometrics. 12 (4): 261-278.

KMA (Korea Meteorological Administration). Climate data of South Korea. 2011. Available: Http://www.kma.go.kr. (May 2011).

Lawrence, J. 1994. Introduction to Neural Networks, California Scientific Software Press. ISBN 1-883157-00-5.

Lee, J. H., M. J. Yu, K. W. Bang, and J. S. Choe. 2003. Evaluation of the Methods for First Flush Analysis in Urban Watersheds. Water Science and Technology. 48 (10): 167-176.

Lee, J., J. Han, H. K. Kumar, J. Choi, H. K. Byeon, J. Choi, J. Kim, M. Jang, H. Park, and K. An. 2011a. Relationship between three major stream assemblages and their environmental factors in multiple spatial scales. International Journal of Limnology. 47: S73-S89.

Lee, S., S. Hwang, J. Lee, D. Jung, Y. Park, and J. Kim. 2011b. Overview and application of the National Aquatic Ecological Monitoring Program (NAEMP) in Korea. International Journal of Limnology. 47: S3-S14.

Lek, S., M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga, and S. Aulagnier. 1996. Application of neural networks to modelling nonlinear relationships in ecology. Ecological Modelling. 90: 39-52.

Li, F., N. Chung, M. Bae, Y. Kwon, and Y. Park. 2012. Relationships between stream macroinvertebrates and environmental variables at multiple spatial scales. Freshwater Biology. 57 (10): 2107-2124.

Longnecker, M. T and R. L. Ott. 2004. A First Course in Statistical Methods. page: 615. Thomson Brooks/Cole.

Mac Nally R. 2002. Multiple regression and inference in ecology and conservation biology: Further comments on identifying important predictor variables. Biodiversity and Conservation. 11 (8): 1397-1401.

Maier, H. R. and G. C. Dandy. 1996. The use of artificial neural networks for the prediction of water quality parameters. Water Resources Research. 32 (4): 1013-1022.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the American Society of Agricultural and Bilogical Engineers. 50 (3): 885-900.

Naiman, R. J., S. E. Bunn, C. Nilsson, G. E. Petts, G. Pinay, and L. C. Thompson. 2002. Legitimizing fluvial ecosystems as users of water: an overview. Environmental Management. 30: 455-467.

Nash, J. E. and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models part I – A discussion of principles. Journal of Hydrology. 10 (3): 282-290.

NIER (National Institute of Environmental Research). 2009. Investigation and evaluation of aquatic ecosystem of South Korean streams and rivers. Final Report. May 2009. Ministry of Environment. S. Korea.

Norgaard, M. 2000. Neural Network Based System Identification Toolbox, Version 2, Department of Automation, Technical Report 00-E-891, Technical University of Denmark, Lungby.

o'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality and Quantity. 41 (5): 673.

Obropta, C. and J. S. Kardos. 2007. Review of urban stormwater quality models: Deterministic, stochastic, and hybrid approaches. Journal of the American Water Resources Association. 43 (6): 1508-1523.

Park, S. H. 1981. Collinearity and Optimal Restrictions on Regression Parameters for Estimating Responses. Technometrics 23 (3): 289-295.

Pearson, K. 1895. Notes on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London. 58: 240-242.

Petts, G. E., M. A. Bickerton, C. Crawford, D. N. Lerner, and D. Evans. 1999. Flow management to sustain groundwater-dominated stream ecosystems. Hydrological processes. 13: 497-513.

Piggott, A. R. and B. P. Neff. 2005. Calculation of streamflow statistics for Ontario and the Great Lakes states. U.S. Geological Survey. Open File Report 2005-1295. Reston, Virginia.

Poff, N. L. and J. V. Ward. 1989. Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Canadian Journal of Fisheries and Aquatic Sciences. 46: 1805-1818.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. C. Stromberg. 1997. The natural flow regime: a paradigm for river conservation and restoration. Bioscience. 47: 769-784.

Power, G., R. S. Brown, and J. G. Imhof. 1999. Groundwater and fish-insights from northern North America. Hydrological Processes. 13: 401-422.

Riad, S., J. Mania, L. Bouchaou, and Y. Najjar. 2004. Rainfall-runoff model using an artificial neural network approach. Mathematical and computer modeling. 40 (7-8): 839-846.

Riseng, C., M.J. Wiley, and R.J. Stevenson. 2004. Hydrologic disturbance and nutrient effects on the structure of benthic stream communities: A covariance structure analysis. Journal of the North American Benthological Society. 23 (2): 309-326.

Riseng, C. R., M. J. Wiley, B. E. Pijanowski, P. W. Seelbach, and R. J. Stevenson. 2010. An ecological assessment of Great Lakes tributaries in the Michigan Peninsulas. Journal of Great Lakes Research. 36: 505-519.

Riseng, C.R., M.J. Wiley, R.W. Black, & M.D. Munn. 2011. Impacts of agricultural land use on biological integrity: a causal analysis. Ecological Applications. 21 (8): 3128-3146.

Sarangi, A. and A. K. Bhattacharya. 2005. Comparison of artificial neural network and regression models for sediment loss prediction from Banha watershed in India. Agricultural Water Management. 78 (3): 195-208.

Seelbach, P. W., L. C. Hinz, M. J. Wiley, and A. R. Cooper. 2011. Use of multiple linear regression to estimate flow regimes for all rivers across Illinois, Michigan, and Wisconsin. Michigan Department of Natural Resources (DNR). Fisheries Division. Fisheries Research Report 2095. Ann Arbor, Michigan.

Smakhtin, V. U. 2001. Low flow hydrology: a review. Journal of hydrology 240: 147-186.

Sousa, S. I. V., F. G. Martins, M. C. M. Alvim-Ferraz, and M. C. Pereira. 2007. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environmental Modelling and Software. 22 (1): 97-103.

SPSS, Inc. 2009. SPSS for Windows Version 12.0. SPSS, Inc., Chicago, Illinois.

Stauffer, J. C., R. M. Goldstein, and R. M. Newman. 2000. Relationship of wooded riparian zones and runoff potential to fish community composition in agricultural streams. Canadian Journal of Fisheries and Aquatic Sciences. 57: 307-316.

Steen, P. J., D. R. Passino-Reader, and M. J. Wiley. 2006. Modeling Brook Trout Presence and Absence from Landscape Variables Using Four Different Analytical Methods. In: Landscape influences on stream habitats and biological communities. American Fisheries Society Symposium 48, 2006. Bethesda, Maryland.

Stevenson, R. J., S. T. Rier, C. M. Riseng, R. E. Schultz, and M. J. Wiley, 2006. Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria. Hydrobiologia 561: 149–165.

Tayfur, G., D. Swiatek, A. Wita, and V. P. Singh. 2005. Case study: finite element method and artificial neural network models for flow through Jeziorsko Earthfill dam in Poland. Journal of Hydraulic Engineering. 131 (6): 431-440.

Trush, W. J., S. M. McBain, and L. B. Leopold. 2000. Attributes of an alluvial river and their relation to water policy and management. Proceedings of the National Academy of Sciences. 97: 11858-11863.

U.S. ACE (U.S. Army Corps of Engineers). 2011. Hydrologic Engineering Center (HEC)-DSSVue 2. 0. 1. Available: Http://www.hec.usace.army.mil/software/hec-dssvue/.

Vayssieres, M., R. E. Plant, and B. H. Allen-Diaz. 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. Journal of Vegetation Science. 11: 679-694.

Velleman, P. F. 1997. DataDesk Version 6.0: Statistics Guide. Data Description. Ithaka, NY.

WAMIS (Water Management Information System). 2011. The Ministry of Land, Infrastructure, and Transport. Stream flow gauges and flow data. Republic of Korea. Available: Http://wamis.go.kr.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 1997. Landscape-based models that predict July thermal characteristics of Lower Michigan Rivers. Michigan Department of Natural Resources (DNR). Fisheries Division. Fisheries Research Report 2037. Ann Arbor, MI.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 2006. Influence of landscape features on summer water temperatures in lower Michigan streams. Pages 113-128 In R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Influences of landscapes on stream habitats and biotic assemblages. American Fisheries Society. Symposium 48. Bethesda, Maryland.

Wen, C. G. and C. S. Lee. 1998. A neural network approach to multiobjective optimization for water quality management in a river basin. Water Resources Research. 34 (3): 427-436.

Wiley, M. J., S. L. Kohler, and P. W. Seelbach. 1997. Reconciling landscape and site based views of stream communities. Freshwater Biology. 37: 133-148.

Wiley, M. J., P. W. Seelbach, K. E. Wehrly, and J. S. Martin. 2003. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. Pages 201-220 in T. P. Simon, editor. Biological response signatures: indicator patterns using aquatic communities. CRC Press. Boca Raton, Florida.

Wiley M. J., D. W. Hyndman, B. C. Pijanowski, A. D. Kendall, C. Riseng, E. S. Rutherford, S.T. Cheng, M.L. Carlson, J.A. Tyler, R.J. Stevenson, P.J. Steen, P.L. Richards, P.W. Seelbach, J.M. Koches & R.R. Rediske. 2010. A multi-modeling approach to evaluating climate and land use change impacts in a Great Lakes River Basin. Hydrobiologia. 657: 243-262.

Yoon, J., J. Kim, M. Byeon, H. Yang, J. Park, J. Shim, H. Song, H. Yang, and M. Jang. 2011. Relationship between three major stream assemblages and their environmental factors in multiple spatial scales. International Journal of Limnology. 47: S91-S105.

Zar, J. H. 1999. Biostatistical analysis. 4th edition. Prentice Hall, Upper Saddle River, New Jersey.

Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2002. Distributions of stream fishes and their relationship to stream size and hydrology in Michigan's Lower Peninsula. Transactions of the American Fisheries Society. 131: 70-85.

Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2004. Utility of species-specific, multiple linear regression models for prediction of fish assemblages in rivers of Michigan's Lower Peninsula. Michigan Department of Natural Resources, Fisheries Research Report 2072, Ann Arbor.

Zorn, T. G., P. W. Seelbach, E. S. Rutherford, T. C. Wills, S. Cheng, and M. J. Wiley. 2008. A regional-scale habitat suitability model to assess the effects of flow reduction on fish assemblages in Michigan streams. Michigan Department of Natural Resources, Fisheries Research Report 2089, Ann Arbor.

Chapter 4 : Biases arising from sampling gear in stream bioassessments: electrofishing versus cast-netting

Abstract

Since different fishing gears are generally recognized to have different capture efficiencies, with respect to both species and habitats, direct comparison of international datasets on community responses to anthropogenic stressors is potentially fraught with sampling biases. In this study I examined the comparability of stream fish data collected using two different sampling methods commonly employed in North American and in East Asian assessment studies: DC electrofishing unit (EF), and hand cast-netting (CN). Paired sampling with these two gears was conducted at 21 stream sites in the Huron River Watershed, MI, USA, and the catches statistically compared. In general, EF had significantly higher ($p<0.05$) sampling performance for count-based metrics of species richness and relative abundance. However, for the metrics reflecting relative percentages there generally are no significant differences between the two types of gear. When fish sampling gear efficiencies (FSGEs) were analyzed for common metrics based on absolute richness (combined datasets from each gear sampling), EF (89.8 ~ 97.0%) had consistently higher FSGEs than CN (72.2 ~ 81.1%). FSGEs of CN were notably lower for Esociformes, Centrarchidae (Perciformes), and Petromyzontiformes. When these groups were excluded from the comparison, EF and CN based metrics showed no significant difference for many common metrics. Intolerant species counts and percentage of insectivorous individuals, however, continued to show significant gear-related differences. I conclude that any direct comparisons of metrics based on EF and CN methods will require some kind of standardization:

either by a restriction of taxa and metrics employed, or by numerical/statistical calibration of metric data to compensate for inherent collection biases.

Introduction

Various fishing gears and methods exist to sample fish assemblages in wadeable streams and rivers (Hellawell 1978, Sutherland 1996, Barbour et al. 1999, An et al. 2005, Cao et al. 2005). Different gears can be more or less appropriate depending upon specific research goals, operating budgets, field time available, and crew familiarity with gear operation (Diamond et al. 1996, Barbour et al. 1999, Carter and Resh 2001, Bonar and Hubert 2002). As the need for management and conservation of ecosystem health has grown (Hwang et al. 2011, Riseng et al. 2011, Allan et al. 2013), governments and researchers around the world are compiling fish community assemblage data to assess anthropogenic impacts. These studies employ gear and procedures that seem appropriate to local needs, and typically these are used consistently from site to site to avoid issues of gear bias (Cao et al. 2005). Regional- and even global- scale assessments and ecological studies can require meta-analyses of, or pooling of, locally designed survey datasets to address issues of interest at broader geographic scales (Wiley et al 2003, Riseng et al. 2010). When these geographically larger scale assessments include samples from different sampling gears, or with similar gear but different faunas, the underlying catch data can reflect methodological variation in capture efficiency and thus bias the metrics derived from them. Biases of course become a concern whenever data from different sources are compared (Intergovernmental Task Force on Monitoring Water Quality 1995, Barbour et al. 1999, Houston et al. 2002).

Many earlier fishery studies have examined the impact of specific gear and environments on capture efficiency and bias. Sampling accuracies among various fish sampling gears or methods were evaluated in many previous papers using electrofishing units, seines, traps, cast nets, electric seines, and rotenone (Freeman 1984, Bayley et al. 1989, Seelbach et al. 1994, Bayley and Austen 2002, An et al. 2005, Cao et al. 2005). Also, sampling effort (i.e. time invested and replication) is another important aspect related to the accuracy of collected fish data (Justus 1994, Paller 1995, Diamond et al. 1996, Barbour et al. 1999, Furse et al. 2006, Polacik et al. 2008). These studies all examined methods in use in a single region or country. However, few studies exist to examine the accuracy and comparability of fish data as influenced by methods in different countries or continents (An et al. 2005, Furse et al. 2006). An et al. (2005) investigated the gear efficiency of electrofishing and cast-netting as a preliminary study of the introduction of western electrofishing gear and methodologies into Korean stream assessment programs. The workers in the context of the EU Water Framework Directive (EU-WFD) have examined existing rapid bioassessment programs and protocols in European countries to provide practical advice and solutions to evaluate aquatic ecosystem health of European streams and rivers (Furse et al. 2006).

Electrofishing is commonly used in rapid bioassessment programs in the North America (MDEQ 1997, Barbour et al. 1999), because of its perceived greater efficiency compared to seining or other net-based methods. Electrofishing allows standardization of catch per unit of effort (time or distance), is less selective (taxonomically), and is useable in a wider variety of habitats than seines or trap nets. Cast-netting is relatively unknown in western countries, but is a popular fishing gear in many East Asian countries and elsewhere. Cast-netting is the standard freshwater fish sampling gear in South Korea. Advantages include its convenience and mobility (An et al. 2005, NIER 2009), as well as its long history of use in artisanal fisheries of the region. Furthermore, cast-netting is relatively inexpensive (to purchase, maintain, and transport), and usually is less damaging to fish leading to lower sampling mortality. However, the comparability of data collected by cast netting has been rarely tested (but see An et al. 2005).

Preliminary to my comparison of North America and Korean assessment datasets (Chapter 6), this study examined two fish sampling methodologies (electrofishing and cast netting) that are commonly used in rapid bioassessment programs in Michigan and South Korea, respectively. My main goal was to examine how the choice of fish sampling gear affects sampling performance, and specifically to analyze what potential biases in assessment metrics could arise from the two sampling gears. My study provides an objective comparison of fish data collected with these two very different sampling gears, and contributes to our knowledge about effective methodologies for ecological assessment of streams and rivers.

Materials and methods

Study site and periods

Fish sampling was conducted at 21 wadeable stream sites of the Huron River, Michigan (Figure 4.1) with two fish sampling gears commonly used in the Michigan and Korean fish monitoring programs: DC electrofishing (EF) and cast net (CN), respectively. The Huron River Watershed, is 136 miles (219 km) long and drains 900 square miles (2,331 km^2) (HRWC 2003). Seven sites were located in the lower main stem of the Huron River Watershed and fourteen sites were located in Mill Creek, a tributary of the lower Huron River. This area included primarily mixed rural land use patterns, directly and indirectly affecting the variety of fish habitat and nutrient condition. In my selection of sampling sites I tried to include variability in environmental factors influencing the performance of either electrofishing or cast-netting including: stream width, water depth, water temperature, stream flow structure (riffles, runs, and pools), stream substrate types, stream discharge volume, and water velocity (Larimore 1961, Bayley 1985, Bayley et al. 1989, Barbour et al. 1999, An et al. 2005; Table 4.1).

I conducted fish sampling during two seasons, reflecting typical seasonal sampling conditions of both Michigan and Korean fish monitoring programs. The first sampling was in the Fall season during October or November, 2010 and the second sampling was in the Spring during June and early July, 2012. Water temperature, stream width, flow depth, substrate types, flow structure (riffle, run, and pool), velocity, percentages of riparian vegetation, stream discharge, sampling time, sampling start and end time, and brief site information was recorded before fish sampling at each site. Percentages of each habitat structure and substrate were estimated by visual examination. Substrate types were classified into six classes: boulder (>256 mm in diameter), cobble (64-256 mm in diameter), pebble and gravel (2-64 mm in diameter), sand (0.06-2 mm in diameter), and clay and silt (<0.06 mm in diameter). Average value of multiple cross-section profiles was used for stream flow depth and velocity.

Sampling distance and time

Each site was sampled by EF and CN methodologies to test the gear-dependent differences in estimates of fish taxa richness and abundance, and other common fish community assessment metrics. The sampling interval between the two gear samplings was typically 20 hours, with a minimum of at least five hours and maximum of one full day, depending on the sampling time and hours of remaining daylight. The order of execution for the two sampling programs was randomly alternated in order to avoid effects of sampling order on fish species diversity and abundance. Sampling reach distance for each site was generally 20 times the stream width, but 50 meters was the minimum for small streams. Actual sampling reach length ranged from 50 to 200 meters and the actual sampling distance was determined to include multiple examples of the most common habitat units (riffles, runs, and pools).

Electrofishing

Electrofishing was by a single-pass sampling (down-stream to up-stream) using either a tow-barge carrying a 240-volt pulse DC electroshocker (Smith-Root model GPP electrofisher), or if the water was too shallow or narrow, a back-pack shocker (Smith-Root LR-24 Electrofisher) with hand-held collecting dip nets (MDEQ 1997, Smith-Root, Inc. 2007). The back-pack shocker was mostly used for smaller streams where wetted width was less than 4 meters and water depth was less than 0.5 meters. Fish sampling was always conducted in an upstream direction and the sampled reach was not blocked by nets. The backpack unit was used at 300 V, 40-60 Hz, and a 7-8 pulse width setting, and a 1-A output was maintained by altering frequency and pulse-width ranges. The electrofishing tow-barge used a 3-phase AC generator with output rectified to DC that produced about 240 V and 3.0 A of current flow.

Cast-netting

A cast net (mesh size: 5×5 mm; net diameter: 6.5 m) in combination with a small hand seine (mesh size: 4×4 mm and net size: 1×1 m) was used following standard South Korean stream health monitoring protocols (NIER 2009). Cast-netting typically works well in sandy, shallow, and stable flow habitats and for surface and midwater-column fish, but is difficult to use in rocky, deep, and turbulent flow habitats. In order to compensate for this weakness a small hand seine is used in addition to the cast net to sample fish in rocky and vegetated habitats. The cast netting team consisted of two persons. One individual with a cast net worked through the entire reach, and the other alternately using both a cast net and a hand seine. The Korean protocol recommends approximately an hour for fish sampling based on new taxa occurance by cast-netting and habitat types covered.

Sampling crews and fish identification

To avoid sampling errors related to sampling proficiency, crews were well experienced in the use of either the EF or CN sampling methods. Cast-netting was conducted only by trained Korean crew members who have recently worked on Korean monitoring projects.

Species names and counts of fish collected were recorded according to the following methods of species identification. Hubbs and Lagler (1964) was used as the primary key for identification of all game fish. Smith (1988) was used for nongame fish, but Hubbs and Lagler (1964) was also used for verification of identification. Vladykov and Kott (1980) was used for additional information on Petromyzonidae (lampreys). After identification, all fish were released immediately. If field identification was unsure or impossible, a few samples of each species from one site were placed preserved with 70% alcohol and labeled for date, site, and species name for later identification.

Fish indicator metrics and data analysis

Fish collections were summarized with twelve commonly used assessment metrics: total number of species (*nTotSp*), total number of individuals (*nTotIn*), number of intolerant species (*nIntSp*), number of tolerant species (*nTolSp*), number of omnivorous species (*nOmnSp*), number of insectivorous species (*nInsSp*), number of piscivorous species (*nPisSp*), percent of intolerant individuals (*pIntIn*), percent of tolerant individuals (*pTolIn*), percent of omnivorous individuals (*pOmnIn*), percent of insectivorous individuals (*pInsIn*), and percent of piscivorous individuals (*pPisIn*; Table 4.2). These fish indicator metrics were chosen to describe representative measures of richness, composition, tolerance guilds, and feeding guilds for a variety of fish assemblages and reflected fish indicator metrics of both MDEQ Procedure-51 (MDEQ 1997) and K-IBIF (NIER 2009).

Fish Sampling Gear Efficiency (FSGE) was estimated for number of species in each order, each family, and six indicator metrics (total number of species, number of intolerant species, number of tolerant species, number of omnivorous species, number of insectivorous species, and number of piscivorous species). FSGE was estimated to evaluate the sampling performance of each sampling gear and was defined as follows,

$$\mathrm{FSGE} = 100 \times (\mathrm{SR_G} / \mathrm{SR_{AD}}) \qquad (1)$$

where SR_G indicates species richness of each data set for each gear and SR_{CB} indicates species richness of the all data from both gear (AD). AD data set was used as a maximum species richness estimate of each sampling site because I could not measure total species list for each site. FSGEs for fish indicator metrics with proportions were not estimated because individual numbers of AD data set can be overlapping due to fish release after sampling and thus overestimated by two sampling gears.

A standardization process was performed for each fish metric separately in order to examine whether differences and biases from sampling gear could be corrected or not. The absolute value of z represents the distance between the raw score and the metric mean in units of standard deviation. The standard score of a raw score χ was

$$z = (\chi - \mu)/\sigma \qquad (2)$$

where μ was the mean of each metric for each sampling gear and σ was the standard deviation of each metric for each sampling gear.

Statistical analysis

Statistical summaries (mean, median, standard deviation, minimum, and maximum), box plots, and scatter plots were conducted using Datadesk 6.0 (Velleman 1997), while paired samples t-test, Pearson correlation, and ANCOVA tests were performed using SPSS 12.0 (SPSS, Inc. 2003).

Results

Characteristics of stream sites and fish habitats

Measured variables potentially influencing performance of the sampling gears (Table 4.1) varied widely across sampling sites. Stream width ranged from 0.65 to 31.5 m with an average of 11.31 m, representing variation from headwater sites to large mainstem (but still wadeable) river sites. Average water depth ranged from 0.25 to 0.75 m with an average of 0.47 m and average velocity ranged from 0.01 to 0.24 m/sec. Water temperature also showed large variation, ranging from 3.6 to 28.4 ºC (average of 18.0 ºC). Particularly, water temperature (n= 10) during spring sampling ranged from 22.4 to 28.4 ºC with an average of 25.5 ºC, whereas water temperature (n= 11) during fall sampling ranged from 3.6 to 15.8 ºC with an average of 11.3 ºC. Mean percentages of channel area as riffle, run, and pool habitats were 23.6, 41.0, and 35.7%, respectively. Percentages of fine (clay, silt, and sand) and coarse (gravel, cobble, and boulder) substrates ranged from 10.0 to 95.0% with average of 58.6% and from 5.0 to 90.0% with average of 41.4%, respectively.

Comparison of sampled fish taxa richness, abundance, and occurrence

Sampled fishes (total species data set: TS) collected by both electrofishing (EF) and cast-netting (CN) included 8 orders, 11 families, 46 species, and 3,590 individuals (Table 4.3). In general, EF captured more fish species and individuals than CN (Table 4.3). Specifically, EF samples captured 7 orders, 10 families, 44 species, and 2,103 individuals, as compared to 7 orders, 10 families, 38 species, and 1,487 individuals from the CN samples. However, Cyprinodontiformes were not captured with the EF method, whereas Petromyzontiformes were not captured with the CN method. *Notropis atherinoides* (emerald shiner) and *Fundulus notatus* (blackstripe topminnow) were not captured with the EF method, whereas *Moxostoma duquesnei* (black redhorse), *Erimyzon oblongus* (creek chubsucker), *Notemigonus crysoleucas* (golden shiner), *Esox lucius* (northern pike), *Pomoxis nigromaculatus* (black crappie), *Lepomis gulosus* (warmouth bass), *Lampetra appendix* (American brook lamprey), and *Ictalurus punctatus* (channel catfish) were not captured with the CN method.

Values of fish metrics based on taxa counts and individual numbers (Table 4.4) were relatively higher in EF than CN samples. However, for fish community metrics based on relative percentages of feeding and tolerance guilds (five metrics), the CN method produced relatively higher values than the EF method, with the exception of the *pOmnIn*. For example, for *pOmnIn* the mean value of CN samples was 26.2% (0.0 to 74.7%), whereas the mean of EF samples was 27.4% (0.0 to 72.5%). Fish metrics representing species richness (five metrics) and individual abundance (one metric) showed significant differences ($p<0.05$) between CN and EF methods, although the *nPisSp* metric differed only marginally ($p= 0.057$) (Table 4.5). All metrics were significantly correlated between sampling gears, again with the exception of the *nPisSp*, which showed no significant correlation ($p>0.05$) between CN and EF samples (Table 4.6). Regression slopes of CN and EF data ranged from 1.0675 (*nTotIn*) to 1.2678 (*nIntSp*), indicating that the EF method had consistently higher values for these metrics (Figure 4.2).

However, fish metrics reflecting relative percentages of the (numerical) catch showed somewhat different results (Table 4.5 and Figure 4.3). Fish metrics for *pTolIn*, *pIntIn*, *pPisIn*, and *pOmnIn* did not significantly differ ($p>0.05$) with sampling gear. The exception was *pInsIn*, where the CN method produced significantly higher percentages of insectivorous individuals than the EF method ($p=0.007$). The Pearson correlations also showed similar trends with strong significant correlations ($p<0.05$) between EF and CN samples except for *pPisIn* (Table 4.6). Regression slopes (zero intercept; Figure 4.3) for these metrics ranged from 0.53 (*pPisIn*) to 0.98 (*pOmnIn*).

Frequency of occurrence

Site occurrence frequencies and relative abundances of fish species were well described by both fish sampling gears (Table 4.3 and Figure 4.4), although there were some small differences due to gear. The EF method found that *Lepomis cyanellus* (green sunfish) and *Hypentelium nigricans* (northern hogsucker) had the highest frequency of occurrence (18 sites) followed by *Etheostoma nigrum* (Johnny darter, 16 sites), *Lepomis macrochirus* (bluegill, 15 sites), *Cottus bairdii* (mottled sculpin, 15 sites), *Semotilus atromaculatus* (creek chub, 15 sites), and *Etheostoma blennioides* (greenside darter, 14 sites), respectively. In contrast the CN method showed that *Etheostoma nigrum* (Johnny darter) had the highest occurrence (16 sites) followed by *Etheostoma blennioides* (greenside darter, 15 sites), *Cottus bairdii* (mottled sculpin, 15 sites), *Semotilus atromaculatus* (creek chub, 15 sites), *Hypentelium nigricans* (northern hogsucker, 15 sites), *Lepomis macrochirus* (bluegill, 13 sites), and *Lepomis cyanellus* (green sunfish, 13 sites), respectively.

Differences in site occurrence frequencies related to sampling gear indicated that sampling efficiencies of the gears varied by fish species (Table 4.7). The EF method showed good efficiency for most species with the EF/AD ratios ranging from 0.85 (*Catostomus commersoni*: white sucker) to 1.00, although the ratio for *Campostoma anomalum* (central stoneroller) was quite low (0.60). *Campostoma anomalum* (1.33) and *Etheostoma blennioides* (greenside darter, 1.07) were the only fish species, for which cast netting was more efficient than electrofishing. For a number of fish species, gears showed equal sampling efficiency, including

Cyprinella spilopterus (spotfin shiner), *Semotilus atromaculatus* (creek chub), *Catostomus commersoni* (white sucker), *Moxostoma erythrurum* (golden redhorse), *Etheostoma caeruleum* (rainbow darter), *Etheostoma nigrum* (Johnny darter), and *Cottus bairdii* (mottled sculpin). In contrast, the CN method had a relatively poor sampling efficiency for certain species, such as *Lampetra appendix* (American brook lamprey, 0.00), *Lepomis gibbosus* (Pumpkinseed sunfish, 0.13), *Esox americanus* (Grass pickerel, 0.43), *Umbra limi* (Central mudminnow, 0.50), and *Pimephales notatus* (Bluntnose minnow, 0.50).

In terms of overall abundance in the sample set, the EF method found that *Catostomus commersoni* (white sucker) was most abundant (239 individuals) followed by *Hypentelium nigricans* (northern hogsucker, 206 individuals), *Cottus bairdii* (mottled sculpin, 205 individuals), *Semotilus atromaculatus* (creek chub, 200 individuals), and *Lepomis cyanellus* (green sunfish, 168 individuals). Using the CN method *Hypentelium nigricans* was most abundant in the samples (212 individuals) followed by *Semotilus atromaculatus* (204 individuals), *Catostomus commersoni* (153 individuals), *Cottus bairdii* (148 individuals), and *Lepomis cyanellus* (108 individuals), respectively.

Comparison of fish sampling gear efficiency (FSGE)

Fish sampling gear efficiency (FSGE; Figure 4.5A) of EF was excellent across all orders with mean ranges from 90.7% (Cypriniformes) to 100.0% (Esociformes, Petromyzontiformes, Scorpaeniformes, and Siluriformes). However, CN FSGEs were relatively lower (means ranging from 45.0% (Esociformes) to 100% (Scorpaeniformes). Three fish orders (Esociformes, Perciformes, and Petromyzontiformes) showed particularly poor performances (45.0%, 79.0%, and 0.0%, respectively) and were significantly different ($p<0.05$) from the FSGEs of the EF method.

For fish families (Figure 4.5B), EF had excellent FSGEs for most families with mean ranging from 88.69% (Cyprinidae) to 100.0% (Cottidae, Esocidae, Ictaluridae, and Petromizontidae). However, the FSGEs for CN were again lower except for Percidae and means ranged from 42.9% (Esociformes) to 100% (Cottidae). Three fish families (Centrarchidae, Esocidae, and Petromyzontidae) again had much lower capture efficiencies (67.1%, 42.9%, and 0.0%).

Fish sampling gear efficiencies for the six fish community metrics showed similar trends in gear performance (Figure 4.6). FSGEs from EF were higher than those from CN. Mean FSGEs for the EF method ranged from 89.8% (*nTotSp*) to 97.0% (*nIntSp*). Mean FSGEs for the CN method ranged from 72.2% (*nIntSp*) to 81.1% (*nInsSp*). Interestingly, *nIntSp* showed the highest FSGE (97.0%) in EF method, whereas it was the lowest (72.2%) in CN method. For *nPisSp*, the FSGEs were 94.4% and 74.1% for EF and CN methods, respectively. However, paired sample T-test showed no significant difference ($p>0.05$) between methods for this metric, whereas all of the other fish metrics differed significantly by gear ($p<0.05$).

Impacts of environmental factors

FSGE of cast-netting was generally influenced by several environmental factors and stream characteristics, whereas electrofishing was relatively insensitive (Table 4.8 and Figure 4.7). Specifically, ANCOVA tests found that water temperature significantly influenced FSGE and led to some biases in *nTotSp*, *nIntSp*, and *nInsSp*, but not for *nTolSp*, *nPisSp*, and *nOmnSp*. Ratio of fine and coarse substrates, ratio of riffles and pools, number of logs, percentages of riparian vegetation, wetted width, mean water depth, and mean velocity did not have significance influence on FSGEs except for number of logs for *nTolSp* (Table 4.8). FSGEs for CN samples were positively correlated with an increase of water temperature, riparian vegetation, mean velocity, percentage of fine substrates, and percentage of riffles. CN efficiencies however were negatively correlated with stream flow width, mean water depth, percentage of coarse substrates, percentage of pools, and number of logs.

Discussion

This study of typical North American and East Asian fish rapid assessment sampling methodologies for stream fishes finds some differences between electrofishing and cast-netting in bioassessment results. I particularly focused on how these gears affected fish data collected, and the potential biases that might be associated with them. The practical advantages and disadvantages of electrofishing are well discussed elsewhere and electrofishing methods have been widely used in many places (Barbour et al. 1999), whereas cast-netting has been less frequently addressed in the scientific literature and seldom examined in terms of accuracy and efficiency (An et al. 2005). I found that cast net sampling was particularly inefficient for Esociformes, Centrarchidae, and Petromyzontiforms (Figure 4.5). However, cast netting also showed slightly better performance for the bottom-associated Percids (consisting here only of darters) than did electrofishing. This is likely related to recovery errors during electrofishing because darter species lack an air bladder and do not typically float to the water surface when electrically stunned.

The differences and biases between fish sampling gear were reflected in values of the twelve fish metrics reported in this study. Of the more than 40 fish indicator metrics currently used in the RBP (Barbour et al. 1999) I examined twelve representing aspects of taxa richness, relative abundance, tolerance guilds, and trophic guilds. I chose metrics here based on the RBP methodology currently used in Michigan and South Korea, because ultimately I am interested in comparing data sets from these two regions. Selection of fish metrics can be crucially important to analyze efficiency (FSGE) of particular sampling gear (Diamond et al. 1996, Barbour et al. 1999, Cao et al. 2005).

I found that metric values reflected efficiency differences between EF and CN samples (Table 4.5), although metric values were for the most part strongly correlated between sampling gears (except *nPisSp*; Table 4.6). Piscivorous fish species sampled in this study included largemouth bass, smallmouth bass, northern pike, grass pickerel, and channel catfish (MDNR 2002). CN sampling efficiency was relatively good for largemouth bass and smallmouth bass, but extremely poor for channel catfish, northern pike, and grass pickerel. Likewise, the analysis of species richness and site occurrence showed that CN had very poor sampling performance for sunfish species, mudminnow, and lamprey.

Electrofishing showed relatively high and stable sampling efficiencies for most orders and families compared to cast-netting (Figure 4.5; 92.5% to 100.0%). However, EF sampling was lower than CN sampling for the Cyprinidae (88.7%) and Percids (92.1%). This pattern reflects differences in the morphology and behavior of these groups. In the locations sampled these fishes were small in size and relatively hard to see and pick up with nets in turbid, turbulent or very shallow water. On the other hand, if cast-netting initially captured small fishes they were very effectively retrieved; thus its sampling efficiency for darters, and small minnows and dace was relatively better.

Since ultimately my interest is in comparing community metric data from South Korea and Michigan, these differences in bias of the two gears are important. Based on the FSGE analysis, species from three fish groups (Esociformes, Centrarchidae, and Petromyzontiforms) were primarily responsible for the poor sampling performance of CN (Figure 4.5). Of these, Esociformes do not occur in South Korean watersheds and there are only two introduced Centrarchidae fish species (bluegill and largemouth bass), both of which are easily caught by cast net (Kim and Park 2002). There are three lamprey species in Korea and so we should not expect these to be represented in the cast net datasets. When the catch data were statistically retested with either these three (Esociformes, Centrarchidae, and Petromyzontiforms) or two (Esociformes and Centrarchidae) groups removed, there was little difference between EF and CN methodologies (Tables 4.6 and 4.9); most of the metrics were not significantly different ($p>0.05$) (except *nIntSp* ($p=0.016$) and *pInsIn* (0.036)). Thus, such an exclusion might be a necessary step before any explicit fish data comparison is made.

A statistically significant influence of several environmental factors was also observed in the FSGEs of CN samples, whereas EF sampling exhibited consistently high performance across the ranges of all factors examined (Table 4.8 and Figure 4.7). Water temperature was one of the most influential environmental factors that had statistically significant influences on FSGEs for certain fish metrics. Particularly, CN showed very poor performance in lower water temperature and this can be interpreted with fish behaviors in lower water temperature. Fish activity is generally reduced with lower temperature and fish tend to move to deeper water and rest in more stable habitats (Diana 2004), where cast net generally cannot catch fish easily. Also, previous studies have reported that turbidity and conductivity influence the sampling performance of EF, but I could not find any significant effects of these factors on FSGEs in my dataset. However, Cyprinidae and Percidae showed relatively poor FSGEs due to their size and shapes leading to poor visibility in water, which suggests the low turbidity and high clarity of water in my sampling areas helped maximize EF performance in this study. Factors which affected CN efficiency were related to the channel habitat types, and included water temperature, flow types, substrate types, stream wetted width and depth, and vegetation. The three fish groups (Esociformes, Centrarchidae, and Petromyzontiforms) with poor FSGEs using CN were, not surprisingly, typically associated with habitats that tended to negatively bias CN performance.

Ecological significances and implications

Fish sampling is often influenced by the availability of labor, budgets, and time. Cast-netting can be a preferred and cost-effective sampling gear in many settings, depending on project purposes and operational constraints. Cast-netting has several advantages including smaller crew numbers, convenience, transportability, and cost for investment and maintenance. Cast-netting overall showed a slightly lower performance than electrofishing and exhibited larger variations in sampling performance for certain fishes. However, comparison of fish metrics indicated that there were strong correlations between estimates provided by the two gears and only small statistical differences in metric values. Thus, I conclude that data based on cast-netting can certainly be effectively applied in ecological monitoring and assessment studies where rapid assessment is an important goal given limitations of time and budget. Furthermore, I conclude that despite some differences in efficiency and bias, comparisons of electrofishing based and cast-net based data sets are possible. Removal of certain fish taxa from the data may help standardize the data for the few strong biases I detected here (e.g. lampreys). Further numerical calibration of metric data may also prove useful since count and richness data seem roughly proportional if not always equivalent

Table 4.1. Summary statistics of environmental factors that may have significant influences on fish sampling performance with electrofishing and cast-netting sampling gears. SD indicates standard deviation.

Stream variables (n= 21)	Mean	Median	SD	Min	Max
Wetted width (m)	11.31	7.45	9.84	0.65	31.5
Average water depth (m)	0.47	0.50	0.16	0.25	0.75
Average velocity (cms)	0.14	0.16	0.07	0.01	0.24
Water temperature (°C)	18.0	15.8	8.1	3.6	28.4
Percentage of riffles	23.6	20.0	16.4	0.0	70.0
Percentage of runs	41.0	40.0	11.8	20.0	60.0
Percentage of pools	35.7	30.0	18.2	10.0	80.0
Percentage of clay (CL)	11.2	0.10	0.10	0.00	0.30
Percentage of silt (SI)	16.2	10.0	16.9	0.00	55.0
Percentage of sand (SA)	31.2	30.0	13.9	10.0	60.0
Percentage of gravels (GR)	20.7	20.0	14.7	5.0	90.0
Percentage of cobbles (CO)	13.6	10.0	12.7	0.0	40.0
Percentage of Boulders (BO)	7.1	5.0	11.6	0.0	40.0
Percentage of fine substrates (CL, SI, SA)	58.6	70.0	30.8	10.0	95.0
Percentage of coarse substrates (GR, CO, BO)	41.4	30.0	30.8	5.0	90.0
Percentage of riparian vegetation	31.1	25.0	22.8	5.0	95.0

Table 4.2. Fish indicator metrics used for the Korean Index of Biological Integrity using Fish (K-IBIF), the South Korean National River Assessment Program (NIER 2009), the biological monitoring program of the Michigan Department of Environmental Quality (MDEQ Procedure-51) (MDEQ 2002), and this study.

	K-IBIF	**MDEQ Procedure-51**	**This study**
Metric 1	Number of endemic species	Total number of fish species	Total number of fish species (*nTotSp*)
Metric 2	Number of riffle benthic species	Number of darter species	Number of tolerant species (*nTolSp*)
Metric 3	Number of sensitive species	Number of sunfish species	Number of intolerant species (*nIntSp*)
Metric 4	Proportion of tolerant individuals	Number of sucker species	Number of picivorous species (*nPicSp*)
Metric 5	Proportion of omnivorous individuals	Number of intolerant species	Number of insectivorous species (*nInsSp*)
Metric 6	Proportion of insectivorous individuals	Percentage of total sample as omnivores	Number of omnivorous species (*nOmnSp*)
Metric 7	Individual numbers of endemic species	Percentage of total sample as insectivorous fish	Total number of fish individuals (*nTotIn*)
Metric 8	Proportion of abnormal individuals	Percentage of total sample as piscivores	Percentage of total sample as tolerant species (*pTolIn*)
Metric 9		Percentage of total sample as tolerant species	Percentage of total sample as intolerant species (*pIntIn*)
Metric 10		Percentage of total sample as simple lithophilic spawners	Percentage of total sample as piscivores (*pPisIn*)
Metric 11			Percentage of total sample as insectivores (*pInsIn*)
Metric 12			Percentage of total sample as omnivores (*pOmnIn*)

Table 4.3. Fish species composition, individual numbers collected, and site occurrence frequency of each species for cast-netting (CN), electrofishing (EF), and all data from both gear (AD).

Order	Family	Species	Total individual numbers collected			Site occurrence frequency		
			CN	EF	AD	CN	EF	AD
Cypriniformes	Catostomidae	Black redhorse	0	1	1	0	1	1
		Creek Chubsucker	0	1	1	0	1	1
		Golden redhorse	9	29	38	4	4	4
		Lake Chubsucker	1	1	2	1	1	1
		Northern hogsucker	212	206	418	15	18	19
		White sucker	153	239	392	11	11	13
	Cyprinidae	Blacknose dace	50	44	94	3	5	5
		Bluntnose minnow	23	34	57	3	6	6
		Central stoneroller	4	7	11	4	3	5
		Common carp	2	5	7	2	3	3
		Common shiner	25	21	46	4	5	5
		Creek Chub	204	200	404	15	15	16
		Emerald shiner	1	0	1	1	0	1
		Fathead minnow	3	11	14	1	1	2
		Golden shiner	0	1	1	0	1	1
		Hornyhead chubs	5	4	9	2	2	3
		River chub	2	7	9	1	3	3
		Roseyface shiner	12	9	21	1	2	2
		Sand Shiner	19	13	32	3	2	3
		Silverside shiner	3	5	8	1	2	2
		Spotfin shiner	32	23	55	4	4	4
		spottail shiner	29	5	34	1	1	1
Cyprinodontiformes	Fundulidae	Blackstripe topminnow	1	0	1	1	0	1

Table 4.3. Continued.

Order	Family	Species	Total individual numbers collected			Site occurrence frequency		
			CN	EF	AD	CN	EF	AD
Esociformes	Esocidae	Grass Pickerel	4	10	14	3	7	7
		Northern pike	0	1	1	0	1	1
	Umbridae	Central mudminnow	36	48	84	3	6	6
Gasterosteiformes	Gasterosteidae	Brook stickleback	15	35	50	1	1	1
Perciformes	Centrarchidae	Black crappie	0	2	2	0	1	1
		Bluegill	57	135	192	13	15	16
		Green sunfish	108	168	276	13	18	19
		Largemouth bass	16	46	62	8	9	10
		PumpkinSeed sunfish	4	25	29	1	8	8
		Rock bass	11	136	147	7	11	11
		Smallmouth bass	63	51	114	7	8	8
		Warmouth Bass	0	1	1	0	1	1
	Percidae	Blackside darter	2	3	5	2	2	2
		Fantail darter	2	5	7	1	1	1
		Greenside darter	87	151	238	15	14	15
		Johnny darter	94	112	206	16	16	17
		Logperch	6	1	7	1	1	1
		Rainbow darter	35	60	95	7	7	7
Petromyzontiformes	Petromyzontidae	American brook lamprey	0	22	22	0	7	7
Scorpaeniformes	Cottidae	Mottled sculpin	148	205	353	15	15	15
Siluriformes	Ictaluridae	Channel catfish	0	1	1	0	1	1
		Stonecat madtom	8	17	25	4	5	5
		Yellow bullhead	1	2	3	1	2	2
Total			1487	2103	3590	196	248	264

Table 4.4. Summary statistics of fish species for all fish indicator metrics used for fish species data comparability (n = 21).

	Electrofishing (EF)					Cast-netting (CN)					All data from both gear (AD)				
Variables	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Total number of species	11.8	11.0	4.1	6.0	24.0	9.3	9.0	3.2	5.0	17.0	12.6	12.0	4.3	6.0	26.0
Total number of individuals	98.6	95.0	50.9	28.0	195.0	70.8	53.0	56.5	21.0	262.0	169.4	148.0	95.9	49.0	457.0
Number of tolerant species	3.8	4.0	1.6	1.0	6.0	3.1	3.0	1.4	1.0	6.0	4.1	4.0	1.5	2.0	6.0
Number of intolerant species	3.8	3.0	2.3	0.0	9.0	2.8	2.0	1.9	0.0	7.0	3.9	3.0	2.4	0.0	10.0
Number of piscivorous species	1.2	1.0	0.9	0.0	3.0	0.9	1.0	0.7	0.0	2.0	1.3	1.0	0.9	0.0	3.0
Number of insectivorous species	7.0	6.0	3.0	2.0	13.0	6.1	5.0	2.9	2.0	12.0	7.5	7.0	3.2	2.0	15.0
Number of omnivorous species	2.4	2.0	1.5	0.0	5.0	1.9	2.0	1.2	0.0	5.0	2.6	2.0	1.4	0.0	5.0
Percentage of tolerant individuals	36.6	33.3	22.4	3.5	91.2	38.5	40.0	26.1	3.4	95.2	37.5	37.2	23.5	3.5	90.2
Percentage of intolerant individuals	35.7	35.3	22.3	0.0	79.7	36.4	45.2	24.3	0.0	81.6	36.6	39.4	22.0	0.0	80.7
Percentage of piscivorous individuals	5.9	5.0	6.1	0.0	26.1	7.0	5.3	7.7	0.0	25.9	5.9	4.9	5.2	0.0	17.7
Percentage of insectivorous individuals	58.5	57.5	17.8	26.1	84.9	65.5	70.0	18.5	25.3	89.9	61.5	61.7	17.3	26.4	86.2
Percentage of omnivorous individuals	27.4	27.6	22.3	0.0	72.5	26.2	22.6	22.2	0.0	74.7	27.0	29.1	21.8	0.0	73.6

Table 4.5. Summary statistics of paired samples T-tests among three data sets collected by different sampling gears. CI and df indicate confidence interval and degree of freedom, respectively. *=0.05; **=0.001.

Fish metrics	Paired differences					t	df	Sig. (2-tailed)
	Mean	Standard deviation	Standard error mean	95% CI of the difference				
				Lower	Upper			
Cast-netting (CN) vs Electrofishing (EF)								
Total number of species	-2.476	2.316	0.505	-3.530	-1.422	-4.900	20	.000**
Total number of individuals	-27.762	48.797	10.648	-49.974	-5.550	-2.607	20	.017*=
Number of tolerant species	-0.667	1.065	0.232	-1.151	-0.182	-2.870	20	.009**
Number of intolerant species	-1.000	1.095	0.239	-1.499	-0.501	-4.183	20	.000**
Number of piscivorous species	-0.381	0.865	0.189	-0.775	0.013	-2.019	20	.057==
Number of insectivorous species	-0.952	1.244	0.271	-1.519	-0.386	-3.508	20	.002**
Number of omnivorous species	-0.524	0.981	0.214	-0.970	-0.077	-2.447	20	.024*=
Percentage of tolerant individuals	1.810	11.513	2.512	-3.430	7.051	0.721	20	.480*=
Percentage of intolerant individuals	0.717	14.005	3.056	-5.658	7.092	0.235	20	.817==
Percentage of piscivorous individuals	1.098	7.992	1.744	-2.540	4.736	0.630	20	.536==
Percentage of insectivorous individuals	6.938	10.667	2.328	2.083	11.794	2.981	20	.007**
Percentage of omnivorous individuals	-1.105	10.871	2.372	-6.053	3.844	-0.466	20	.647==
Cast-netting (CN) vs All data from both gear (AD)								
Total number of species	-3.238	2.189	0.478	-4.234	-22.242	-6.780	20	.000**
Total number of individuals	-98.571	50.895	11.106	-121.738	-75.404	-8.875	20	.000**
Number of tolerant species	-0.952	0.805	0.176	-1.319	-0.586	-5.423	20	.000**
Number of intolerant species	-1.095	1.179	0.257	-1.632	-0.558	-4.256	20	.000**
Number of piscivorous species	-3.190	1.750	0.382	-3.987	-2.394	-8.355	20	.000**
Number of insectivorous species	-1.381	0.973	0.212	-1.824	-0.938	-6.501	20	.000**
Number of omnivorous species	-0.714	0.845	0.184	-1.099	-0.330	-3.873	20	.001**

Table 4.5. Continued.

Fish metrics	Paired differences					t	df	Sig. (2-tailed)
	Mean	Standard deviation	Standard error mean	95% CI of the difference				
				Lower	Upper			
Cast-netting (CN) vs All data from both gear (AD)								
Percentage of tolerant individuals	0.957	6.396	1.396	-1.954	3.868	0.686	20	.501==
Percentage of intolerant individuals	-0.214	8.155	1.780	-3.926	3.498	-0.120	20	.906==
Percentage of piscivorous individuals	1.045	4.465	0.974	-0.987	3.078	1.073	20	.296==
Percentage of insectivorous individuals	3.969	6.302	1.375	1.100	6.837	2.886	20	.009**
Percentage of omnivorous individuals	-0.788	6.659	1.453	-3.819	2.242	-0.543	20	.593==
Electrofishing (EF) vs All data from both gear (AD)								
Total number of species	-0.762	1.044	0.228	-1.237	-0.287	-3.344	20	.003**
Total number of individuals	-70.810	56.523	12.334	-96.539	-45.081	-5.741	20	.000**
Number of tolerant species	-0.286	0.717	0.156	-0.612	0.041	-1.826	20	.083==
Number of intolerant species	-0.095	0.301	0.066	-0.232	0.042	-1.451	20	.162==
Number of piscivorous species	-0.048	0.218	0.048	-0.147	0.052	-1.000	20	.329==
Number of insectivorous species	-0.429	0.811	0.177	-0.798	-0.060	-2.423	20	.007**
Number of omnivorous species	-0.190	0.402	0.088	-0.374	-0.007	-2.169	20	.042*=
Percentage of tolerant individuals	-0.853	5.457	1.191	-3.337	1.631	-0.717	20	.482==
Percentage of intolerant individuals	-0.931	6.163	1.345	-3.736	1.874	-0.692	20	.497==
Percentage of piscivorous individuals	-0.053	3.739	0.816	-1.755	1.649	-0.065	20	.949==
Percentage of insectivorous individuals	-2.969	4.951	1.080	-5.223	-0.716	-2.749	20	.012**
Percentage of omnivorous individuals	0.316	4.659	1.017	-1.805	2.437	0.311	20	.759==

Table 4.6. Two tailed Pearson correlation tests between fish sampling data from different gears for each fish metric. Three fish groups removed were Centrachidae, Esociformes, and Petromyzontiformes. Bold indicates significance at p≤0.05, and bold and italics indicate significance at p≤0.01. "a" indicates no enough data to compare.

Fish metrics	All data			Without three fish groups		
	CN vs EF	CN vs AD	EF vs AD	CN vs EF	CN vs AD	EF vs AD
Total number of species	***.822***	***.872***	***.971***	***.862***	***.906***	***.973***
Total number of individuals	***.592***	***.904***	***.880***	**.521**	***.879***	***.865***
Number of tolerant species	***.740***	***.841***	***.888***	***.630***	***.751***	***.908***
Number of intolerant species	***.883***	***.872***	***.992***	***.893***	***.891***	***.987***
Number of piscivorous species	.405	***.439***	***.970***	a	a	***1.000***
Number of insectivorous species	***.912***	***.953***	***.968***	***.952***	***.960***	***.982***
Number of omnivorous species	***.746***	***.808***	***.962***	***.698***	***.775***	***.954***
Percentage of tolerant individuals	***.899***	***.972***	***.973***	***.749***	***.810***	***.982***
Percentage of intolerant individuals	***.823***	***.943***	***.961***	**.504**	***.704***	***.906***
Percentage of piscivorous individuals	.352	***.831***	***.794***	a	a	***1.000***
Percentage of insectivorous individuals	***.828***	***.940***	***.961***	***.873***	***.922***	***.981***
Percentage of omnivorous individuals	***.880***	***.954***	***.978***	***.876***	***.923***	***.982***

a. The correlation and t could not be computed because the standard error of the difference was 0.

Table 4.7. Comparison of fish species occurrence ratios among cast-netting (CN), electrofishng (EF), and all data from both gear (AD). The fish species that occurred in less than 4 sites in all data from both gear (AD) were removed from the table and the occurrence ratios were sorted from largest to smallest.

CN/AD		EF/AD		CN/EF	
Species	Ratio	Species	Ratio	Species	Ratio
Greenside darter	1.00	Mottled sculpin	1.00	Central stoneroller	1.33
Mottled sculpin	1.00	Rainbow darter	1.00	Greenside darter	1.07
Rainbow darter	1.00	Golden redhorse	1.00	Mottled sculpin	1.00
Golden redhorse	1.00	Spotfin shiner	1.00	Rainbow darter	1.00
Spotfin shiner	1.00	Smallmouth bass	1.00	Golden redhorse	1.00
Johnny darter	0.94	Common shiner	1.00	Spotfin shiner	1.00
Creek Chub	0.94	Stonecat madtom	1.00	Johnny darter	1.00
Smallmouth bass	0.88	Rock bass	1.00	Creek Chub	1.00
White sucker	0.85	Blacknose dace	1.00	White sucker	1.00
Bluegill	0.81	Bluntnose minnow	1.00	Largemouth bass	0.89
Central stoneroller	0.80	Central mudminnow	1.00	Smallmouth bass	0.88
Largemouth bass	0.80	Grass Pickerel	1.00	Bluegill	0.87
Common shiner	0.80	PumpkinSeed sunfish	1.00	Northern hogsucker	0.83
Stonecat madtom	0.80	American brook lamprey	1.00	Common shiner	0.80
Northern hogsucker	0.79	Northern hogsucker	0.95	Stonecat madtom	0.80
Green sunfish	0.68	Green sunfish	0.95	Green sunfish	0.72
Rock bass	0.64	Johnny darter	0.94	Rock bass	0.64
Blacknose dace	0.60	Creek Chub	0.94	Blacknose dace	0.60
Bluntnose minnow	0.50	Bluegill	0.94	Bluntnose minnow	0.50
Central mudminnow	0.50	Greenside darter	0.93	Central mudminnow	0.50
Grass Pickerel	0.43	Largemouth bass	0.90	Grass Pickerel	0.43
PumpkinSeed sunfish	0.13	White sucker	0.85	PumpkinSeed sunfish	0.13
American brook lamprey	0.00	Central stoneroller	0.60	American brook lamprey	0.00

Table 4.8. P values from ANCOVA tests of SGEs for fish metrics. Sampling gear was used as fixed factor and urban and agricultural land uses were used as a covariate.

Covariates	Fish metrics for fish sampling gear efficiency (FSGE)					
	Total number of species	Number of tolerant species	Number of intolerant species	Number of piscivorous species	Number of insectivorous species	Number of omnivorous species
Sampling gear	0.000	0.008	0.000	0.075	0.000	0.109
Water temperature (ºC)	0.003	0.559	0.016	0.645	0.010	0.144
Sampling gear	0.000	0.008	0.000	0.072	0.001	0.114
Ratio of fine and coarse substrates	0.918	0.299	0.606	0.334	0.836	0.316
Sampling gear	0.000	0.008	0.000	0.076	0.001	0.115
Ratio of riffles and pools	0.119	0.359	0.054	0.703	0.400	0.356
Sampling gear	0.000	0.006	0.000	0.075	0.001	0.114
Number of logs	0.567	0.030	0.900	0.663	0.716	0.318
Sampling gear	0.000	0.008	0.000	0.072	0.001	0.118
Percentages of riparian vegetation (%)	0.453	0.622	0.961	0.336	0.537	0.610
Sampling gear	0.000	0.009	0.000	0.075	0.001	0.119
Wetted width (m)	0.343	0.943	0.147	0.573	0.572	0.709
Sampling gear	0.000	0.009	0.000	0.076	0.001	0.119
Mean water depth (m)	0.511	0.943	0.258	0.942	0.767	0.849
Sampling gear	0.000	0.007	0.000	0.076	0.001	0.117
Mean velocity (m/s)	0.266	0.123	0.198	0.857	0.495	0.449

Table 4.9. Summary statistics of paired samples T-tests among three data sets collected by different sampling gears. Fish species of Esociforms, Centrachidae, and Petromyzontidae were excluded in this analysis. CN, EF, and AD indicate cast-netting, electrofishing, and combined data, respectively. CI and df indicate confidence interval and degree of freedom, respectively. *=0.05; **=0.001.

Fish indicators	Paired differences					t	df	Sig. (2-tailed)
	Mean	Standard deviation	Standard error mean	95% CI of the difference				
				Lower	Upper			
Cast-netting (CN) vs Electrofishing (EF)								
Total number of species	-0.762	1.868	0.408	-1.612	0.089	-1.869	20	.076**
Total number of individuals	-11.667	43.557	9.505	-31.494	8.160	-1.227	20	.234*=
Number of tolerant species	-0.286	1.056	0.230	-0.766	0.195	-1.240	20	.229**
Number of intolerant species	-0.429	0.746	0.163	-0.768	-0.089	-2.631	20	.016**
Number of piscivorous species	-0.048	0.218	0.048	-0.147	0.052	-1.000	20	.329==
Number of insectivorous species	-0.381	0.973	0.212	-0.824	0.062	-1.793	20	.088**
Number of omnivorous species	-0.381	0.973	0.212	-0.824	0.062	-1.793	20	.088*=
Percentage of tolerant individuals	-1.817	20.482	4.470	-11.140	7.507	-0.406	20	.689*=
Percentage of intolerant individuals	4.813	25.165	5.491	-6.641	16.268	0.877	20	.391==
Percentage of piscivorous individuals	-0.034	0.156	0.034	-0.105	0.037	-1.000	20	.329==
Percentage of insectivorous individuals	7.322	14.956	3.264	0.514	14.130	2.243	20	.036**
Percentage of omnivorous individuals	-6.694	14.708	3.210	-13.389	0.001	-2.086	20	.050==
Cast-netting (CN) vs All data from both gear (AD)								
Total number of species	-1.381	1.596	0.348	-2.108	-0.654	-3.965	20	.001**
Total number of individuals	-68.238	43.093	9.404	-87.854	-48.622	-7.257	20	.000**
Number of tolerant species	-0.524	0.814	0.178	-0.894	-0.153	-2.950	20	.008**
Number of intolerant species	-0.524	0.814	0.178	-0.894	-0.153	-2.950	20	.008**
Number of piscivorous species	-0.048	0.218	0.048	-0.147	0.052	-1.000	20	.329**
Number of insectivorous species	-0.714	0.956	0.209	-1.150	-0.279	-3.423	20	.003**
Number of omnivorous species	-0.571	0.811	0.177	-0.940	-0.202	-3.230	20	.004**

Table 4.9. Continued.

Fish indicators	Paired differences					t	df	Sig. (2-tailed)
	Mean	Standard deviation	Standard error mean	95% CI of the difference				
				Lower	Upper			
Cast-netting (CN) vs All data from both gear (AD)								
Percentage of tolerant individuals	-1.613	17.437	3.805	-9.550	6.324	-0.424	20	.676==
Percentage of intolerant individuals	2.997	19.315	4.215	-5.795	11.789	0.711	20	.485==
Percentage of piscivorous individuals	-0.018	0.084	0.018	-0.056	0.020	-1.000	20	.329==
Percentage of insectivorous individuals	5.095	11.206	2.445	-0.006	10.196	2.084	20	.050**
Percentage of omnivorous individuals	-4.745	11.097	2.422	-9.796	0.307	-1.959	20	.064==
Electrofishing (EF) vs All data from both gear (AD)								
Total number of species	-0.619	0.865	0.189	-1.013	-0.225	-3.281	20	.004**
Total number of individuals	-56.571	45.938	10.024	-77.482	-35.661	-5.643	20	.000**
Number of tolerant species	-0.238	0.539	0.118	-0.483	0.007	-2.024	20	.056==
Number of intolerant species	-0.095	0.301	0.066	-0.232	0.042	-1.451	20	.162==
Number of piscivorous species								
Number of insectivorous species	-0.333	0.658	0.144	-0.633	-0.034	-2.320	20	.031**
Number of omnivorous species	-0.190	0.402	0.088	-0.374	-0.007	-2.169	20	.042*=
Percentage of tolerant individuals	0.204	5.495	1.199	-2.298	2.705	0.170	20	.867==
Percentage of intolerant individuals	-1.816	9.760	2.130	-6.259	2.626	-0.853	20	.404==
Percentage of piscivorous individuals	0.016	0.072	0.016	-0.017	0.049	1.000	20	.329==
Percentage of insectivorous individuals	-2.227	6.130	1.338	-5.017	0.564	-1.664	20	.112**
Percentage of omnivorous individuals	1.950	5.940	1.296	-0.754	4.654	1.504	20	.148==

.

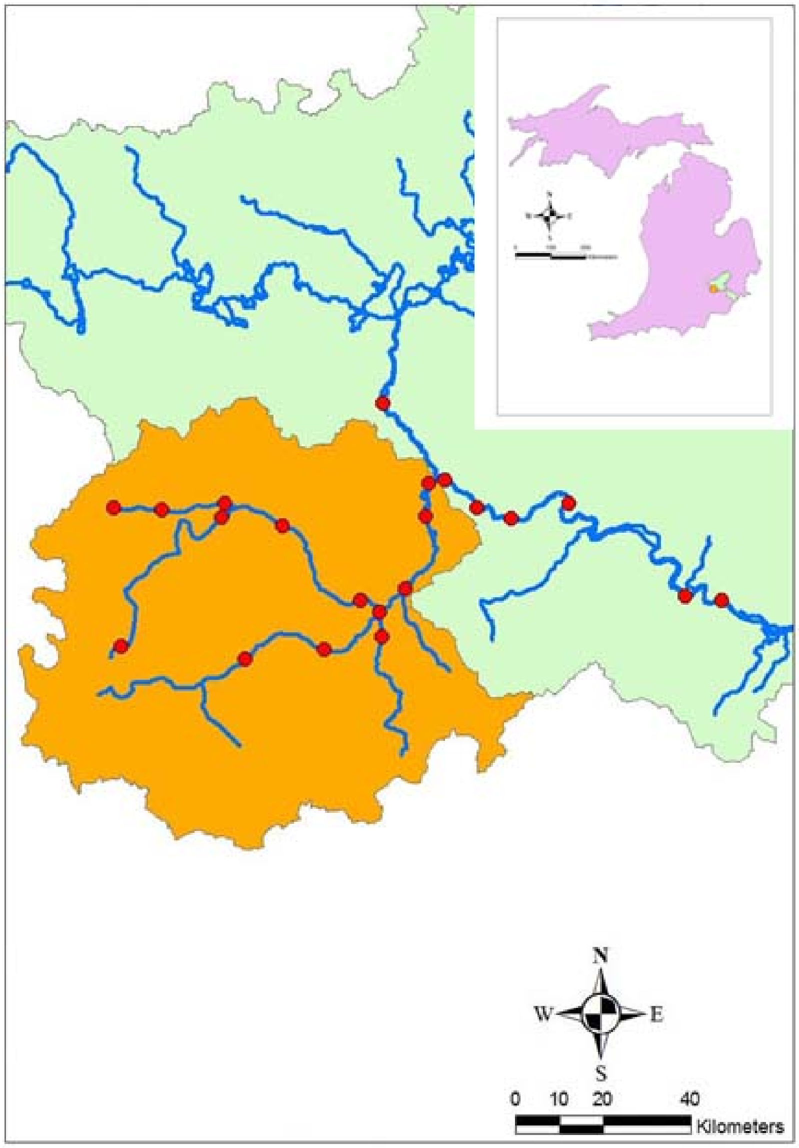


Figure 4.1. Locations of fish sampling sites (n= 21) for electrofishing (EF) and cast-netting (CN).

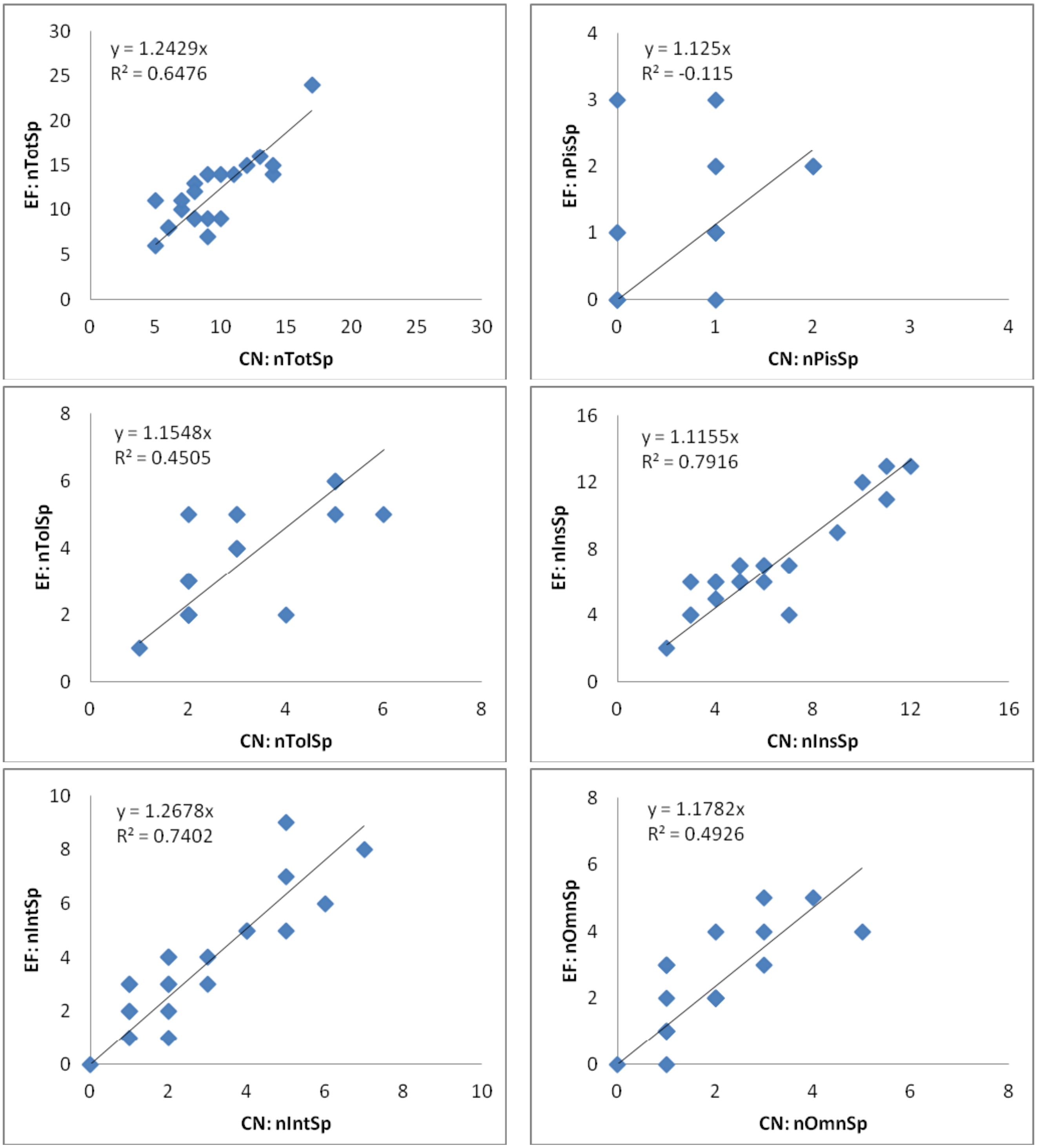


Figure 4.2. Comparison of species numbers between cast-netting (CN) and electrofishing (EF) and correction equations for sampling gear bias.

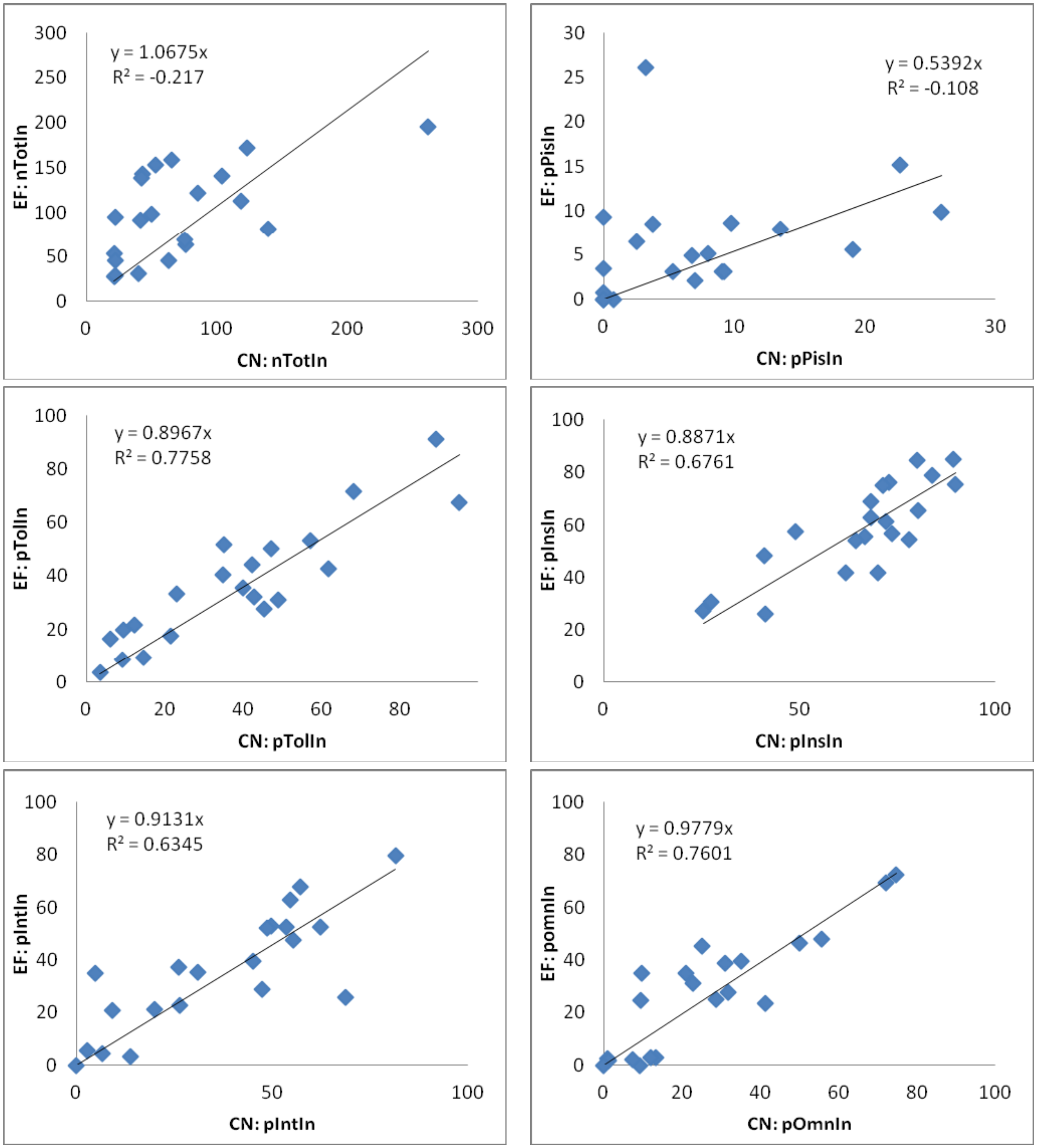
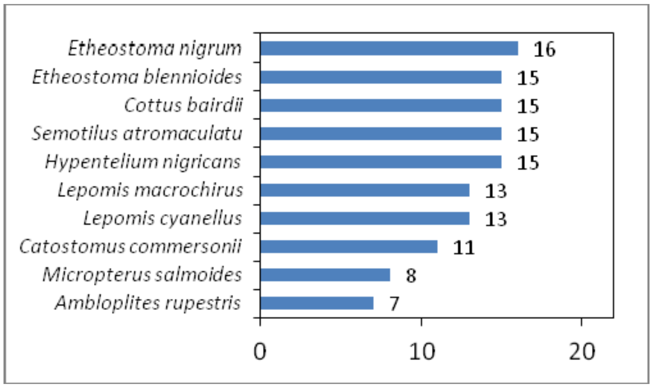
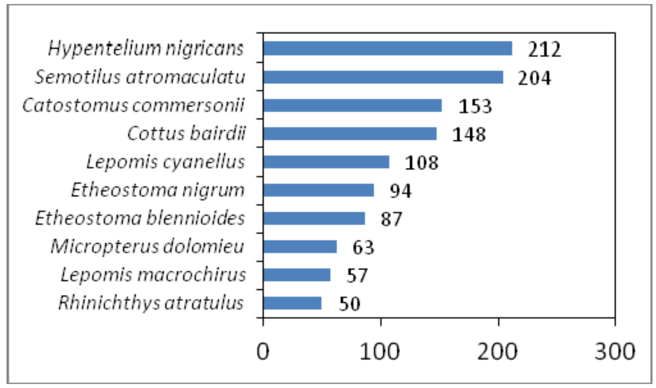


Figure 4.3. Comparison of individual number percentages between cast-netting (CN) and electrofishing (EF) and correction equations for sampling gear bias.

Cast-netting method (CN)

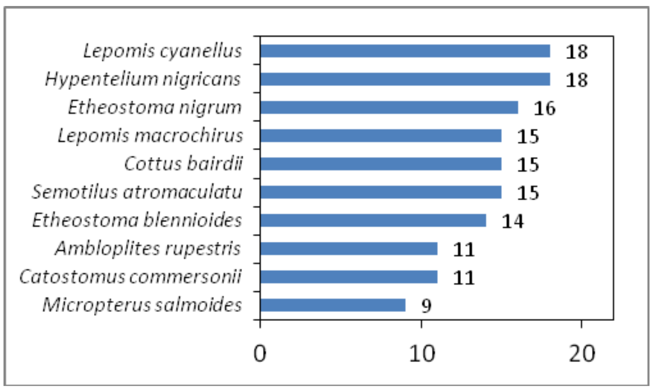


Site occurrence frequency

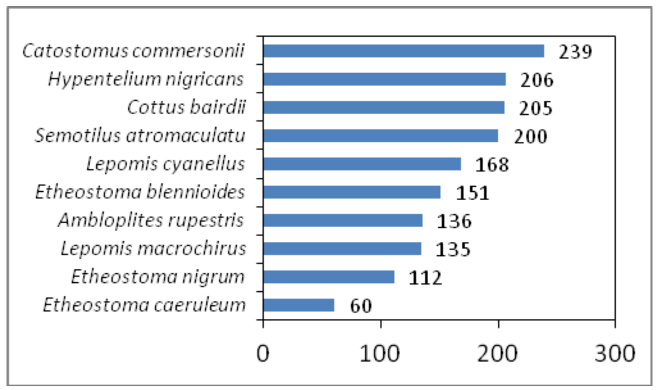


Relative abundance

Electrofishing method (EF)

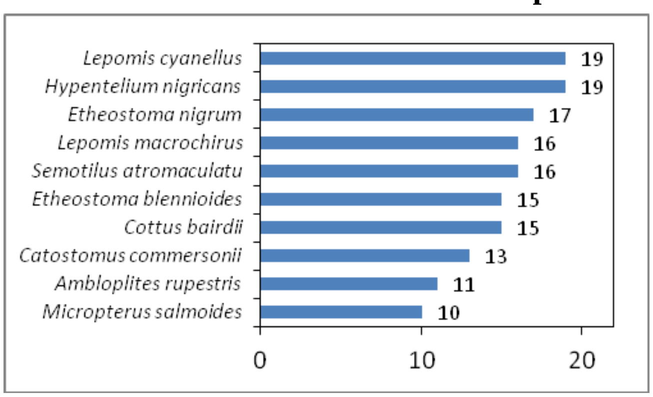


Site occurrence frequency

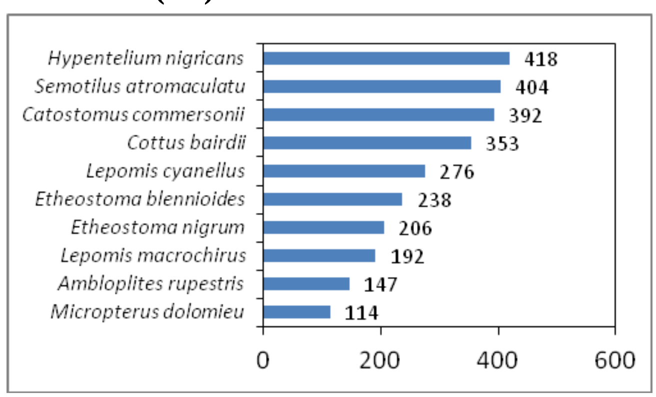


Relative abundance

Total species data combined (TS)



Site occurrence frequency



Relative abundance

Figure 4.4. Site occurrence frequencies and relative fish abundances of the highest top ten fish species for cast-netting method (CN), electrofishing method (EF), and all data from both gear (AD).

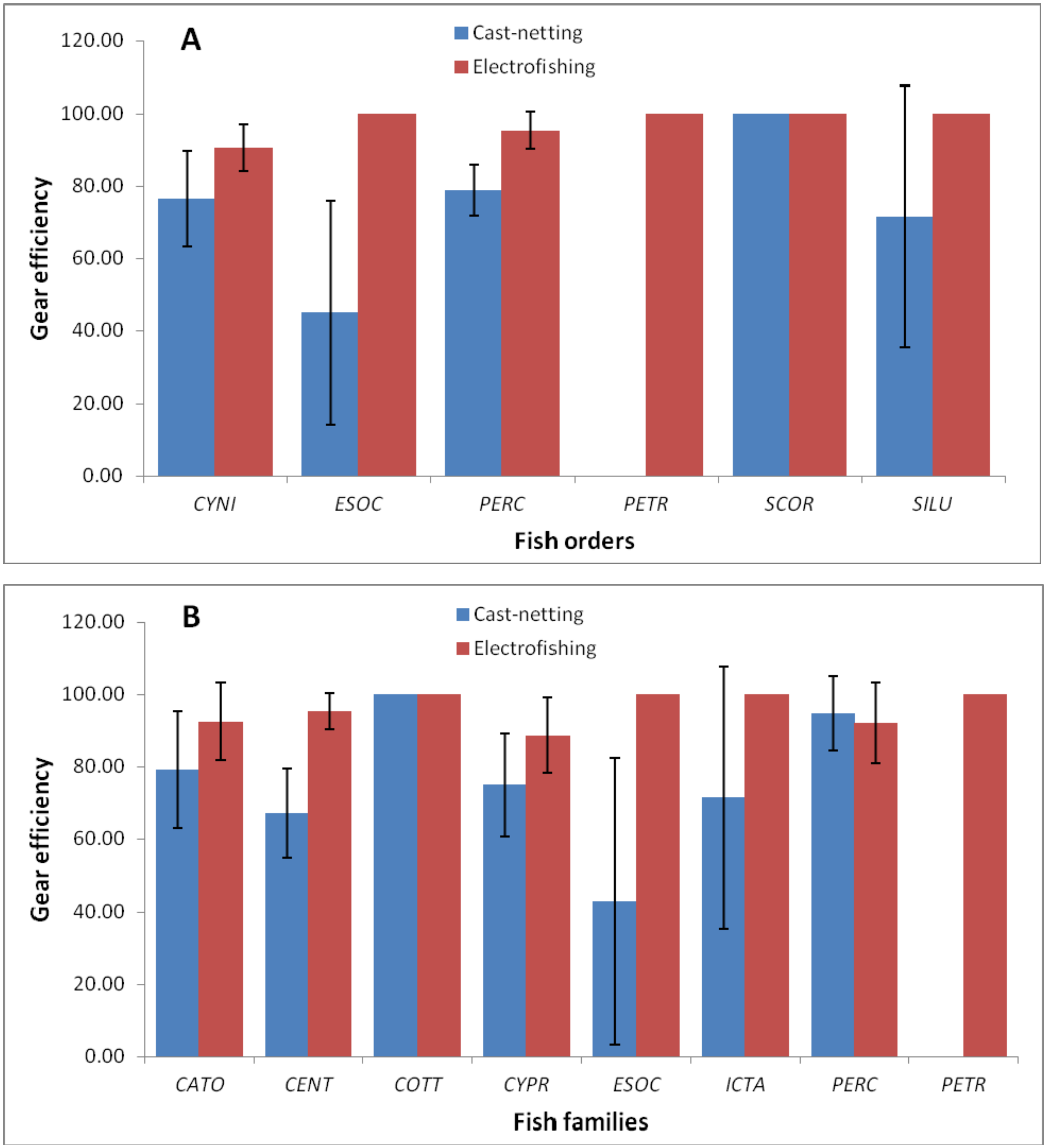


Figure 4.5. Mean fish sampling gear efficiency (SGE) with 95 percentile ranges for cast-netting and electrofishing methods. A) SGE summarized for each order (CYNI: Cypriniformes, ESOC: Esociformes, PERC: Perciformes, PETR: Petromyzontiformes, SCOR: Scorpaeniformes, SILU: Siluriformes). B) SGE summarized for each family (CATO: Catostomidae, CENT: Centrachidae, COTT: Cottidae, CYPR: Cyprinidae, ESOC: Esocidae, ICTA: Ictaluridae, PERC: Percidae, PETR: Petromizontidae). Orders and families collected in less than seven sites were removed from the graphs.

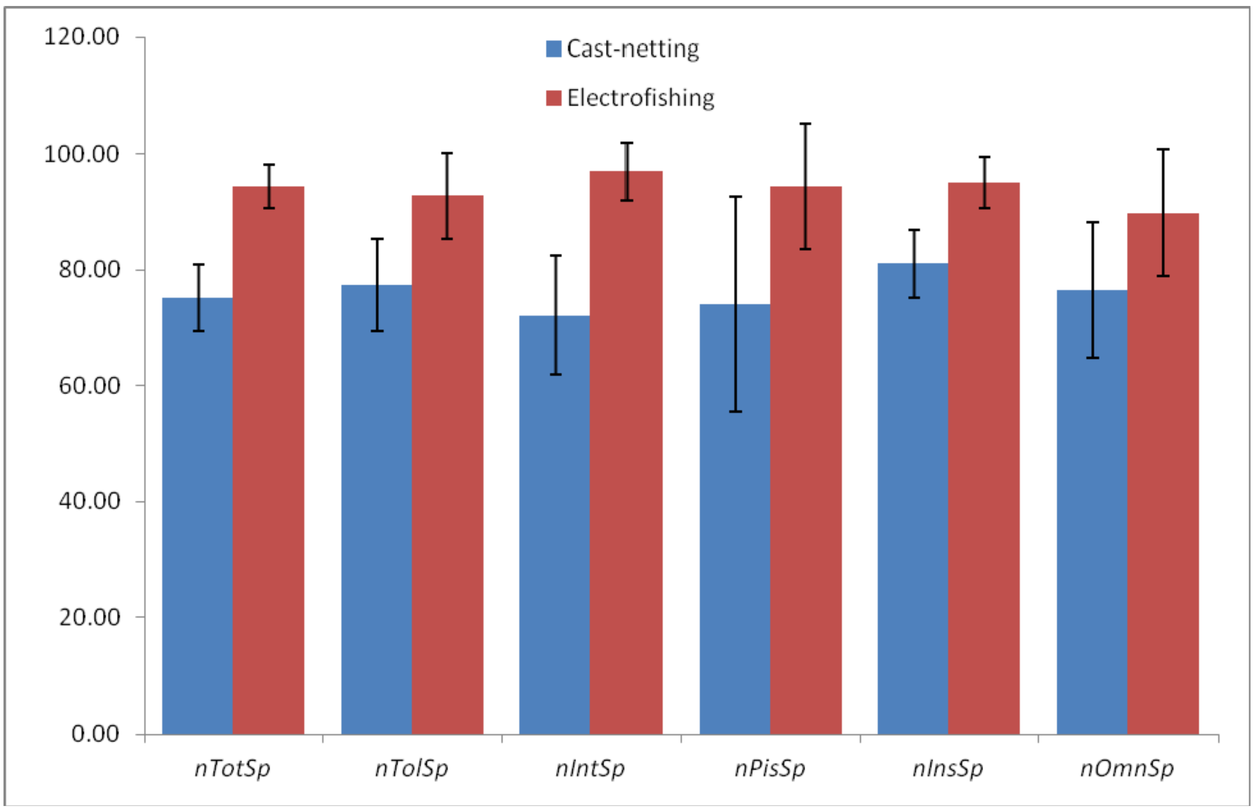


Figure 4.6. Mean fish sampling gear efficiency (SGE) with 95 percentile ranges for cast-netting and electrofishing methods. SGE summarized for each fish indicator metric.

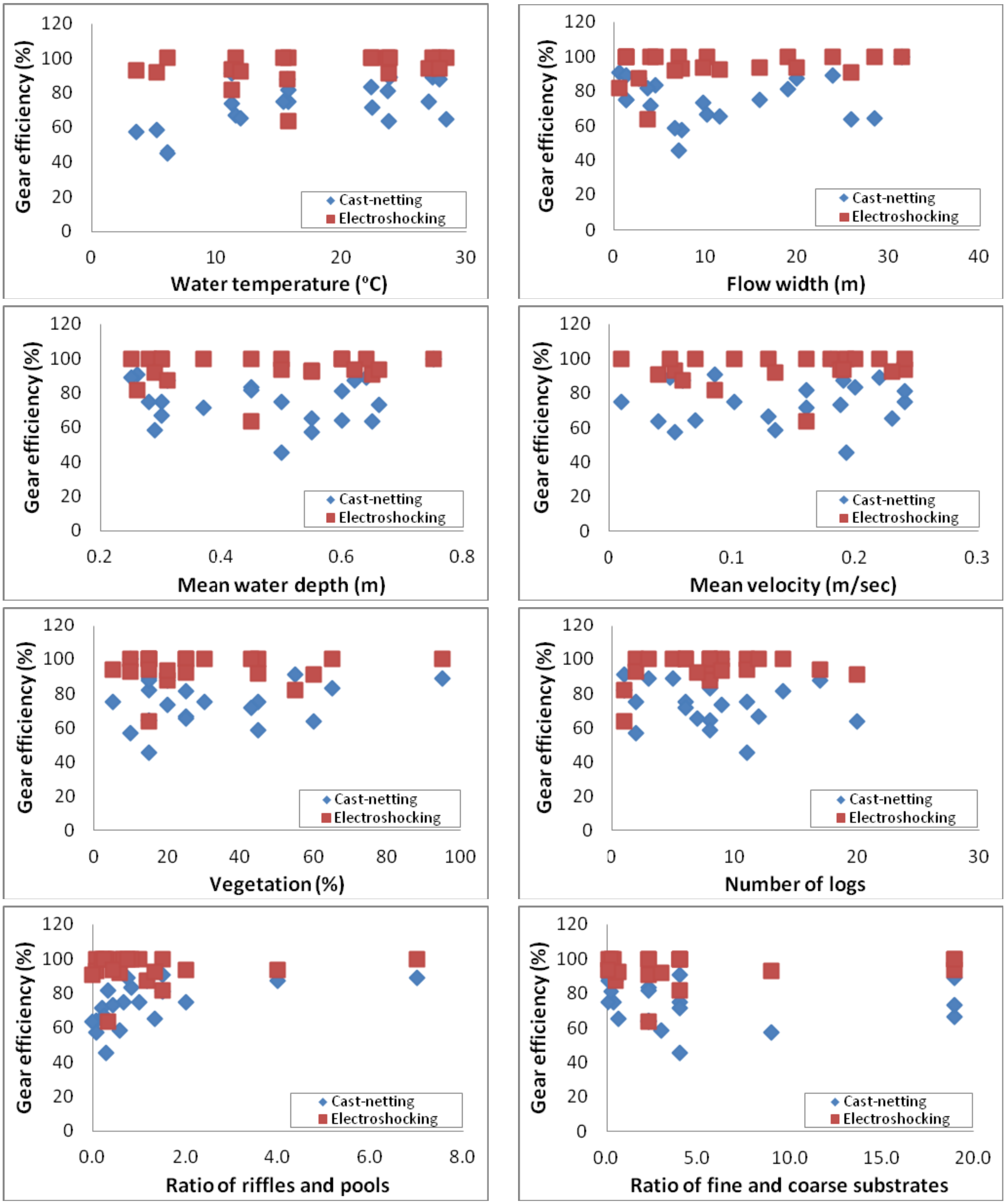


Figure 4.7. Scatter plots showing the relationship between sampling gear efficiency and environmental factors for each sampling gear using total number of species.

Literature cited

Allan, J. D., P. B. McIntyre, S. D. P. Smith, B. S. Halpern, G. L. Boyer, A. Buchsbaum, G. A. Burton Jr., L. M. Campbell, W. L. Chadderton, J. J. H. Ciborowski, P. J. Doran, T. Eder, D. M. Infante, L. B. Johnson, C. A. Joseph, A. L. Marino, A. Prusevich, J. Read, J. B. Rose, E. S. Rutherford, S. P. Sowa, A. D. Steinman. 2013. Joint analysis of stressors and ecosystem services to enhance restoration effectiveness. Proceedings of the National Academy of Sciences. 110 (1): 372-377.

An, K., S. Hwang, D. Bae, J. Lee, and J. Kim. 2005. Introduction of an electrofishing technique for assessments of fish assemblages to Korean watersheds. Korean Journal of Limnology. 38 (4): 482-493.

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

Bayley, P. B. 1985. Sampling problems in freshwater fisheries. Pp: 3-11. In O'hara, K., C. Aprahamian, and R. T. Leah, editors. Proceedings of the forth British freshwater fisheries conference. University of Liverpool, Liverpool, UK.

Bayley, P. B., R. W. Larimore, and D. C. Dowling. 1989. Electric seine as a fish-sampling gear in streams. Transactions of the American Fisheries Society. 118 (4): 447-453.

Bayley, P. B. and D. J. Austen. 2002. Capture efficiency of a boat electrofisher. Transactions of the American Fisheries Society. 131 (3): 435-451.

Bonar, S. A. and W. A. Hubert. 2002. Standard sampling of inland fish: benefits, challenges, and a call for action. Fisheries. 27: 10-16.

Cao, Y., C. P. Hawkins, and A. W. Storey. 2005. A method for measuring the comparability of different sampling methods used in biological surveys: implications for data integration and synthesis. Freshwater Biology. 50: 1105-1115.

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

Diana, J. S. 2004. Biology and ecology of fishes, Second edition. Cooper Publishing Group. Traverse City, MI. pp 20-42 and 399-413.

Diamond, J. M., M. T. Barbour, and J. B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. Journal of North American Benthological Society. 15 (4): 713-727.

Freeman, B. J., H. S. Greening, and J. D. Oliver. 1984. Comparison of three methods for sampling fishes and macroinvertebrates in a vegetated wetland. Faculty Publications and Presentations. Paper 89.

Furse, M., D. Hering, O. Moog, P. Verdonschot, R. K. Johnson, K. Brabec, K. Gritzalis, A. Buffagni, P. Pinto, N. Friberg, J. Murray-Bligh, J. Kokes, R. Alber, P. Usseglio-Polatera, P. Haase, R. Sweeting, B. Bis, K. Szoszkiewicz, H. Soszka, G. Springe, F. Sporka, and I. Krno. 2006. The STAR project: context, objectives, and approaches. Hydrobiologia. 566: 3-29.

Hellawell, J. M. 1978. Biological surveillance of rivers. A biological monitoring handbook. Water Research Centre, Medmenham and Stevenage, England.

Houston, L., M. T., Barbour, D. Lenat, and D. Penrose. 2002. A multi-agency comparison of aquatic macroinvertebrate-based stream bioassessment methodologies. Ecological Indicators. 1: 279-292.

Hubbs, C. L., and K. F. Lagler. 1964. Fishes of the Great Lakes Region. Second Edition. The University of Michigan Press, Ann Arbor, MI. 213 pp.

Huron River Watershed Council (HRWC). 2003. Mill Creek Subwatershed Management Plan, Huron River Watershed (HUC 04090005), Ann Arbor, Michigan.

Hwang, S., S. Lee, and Y. Park. 2011. Ecological monitoring, assessment, and restoration of running waters in Korea. Annales de Limnologie – International Journal of Limnology. 47: S1-S2.

Intergovernmental Task Force on Monitoring Water Quality. 1995. The strategy for improving water – quality monitoring in the United States. Final report of the Intergovernmental Task Force on Monitoring Water Quality. Office of Water Data Coordination, U. S. Geological Survey, Reston, Virginia, Open-File Report 95-742.

Justus, B. 1994. Observations on electrofishing techniques for three catfish species in Mississippi. Proceedings of the annual conference. The Southeastern Association of Fish and Wildlife Agencies. The Mississippi Office of Pollution Control. 48: 524-532.

Kim, I. S. and J. Y. Park. 2002. Freshwater fish of Korea. Kyo-Hak Publishing Co., Ltd., Korea.

Larimore, R. W. 1961. Fish population and electrofishing success in a warm-water stream. Journal of Wildlife Management. 25: 1-12.

Michigan Department of Environmental Quality (MDEQ). 1997. Qualtitative biological and habitat survey protocols for wadeable streams and rivers. Michigan Department Environmental Quality (MDEQ), Surface Water Quality Division, Great Lakes and Environmental Assessment Section, GLEAS Procedure 51 (Revised 2002). Lansing, Michigan.

Michigan Department of Natural Resources (MDNR). 2002. Huron River Plan (Livingston, Washtenaw, Oakland, Monroe and Wayne Counties). Michigan Department of Natural Resources, Fisheries Division. Lansing, Michigan.

NIER (National Institute of Environmental Research). 2009. Investigation and evaluation of aquatic ecosystem of South Korean streams and rivers. Final Report. May 2009. Ministry of Environment. S. Korea.

Paller, M. H. 1995. Relationships among number of fish species sampled, reach length surveyed, and sampling effort in South Carolina coastal plain streams. North American Journal of Fisheries Management. 15 (1): 110-120.

Polacik, M., M. Janac, P. Jurajda, M. Vassilev, and T. Trichkova. 2008. The sampling efficiency of electrofishing for *Neogobius* species in a riprap habitat: a field experiment. Journal of Applied Ichthyology. 24: 601-604.

Riseng, C. M., M. J. Wiley, P. W. Seelbach, and R. J. Stevensen. 2010. An ecological assessment of Great Lakes tributaries in the Michigan Peninsulas. Journal of Great Lakes Research. 36: 505-519.

Riseng, C. M., M. J. Wiley, R. W. Black, and M. D. Munn. 2011. Impacts of agricultural land use on biological integrity: a causal analysis. Ecological Applications. 21 (8): 3128-3146.

Seelbach, P. W., R. N. Lockwood, and J. R. Ryckman. 1994. Efficiency of sampling river fishes with rotenone. Michigan Department of Natural Resources, Mighican, Fisheries Research Report No. 2009.

Smith, G. 1988. Guide to Nongame Fishes of Michigan. University of Michigan, Museum of Zoology.

Smith-Root, Inc. 2007. User's Manual: LR-24 electrofisher. Vancouver, WA.

SPSS, Inc. 2003. SPSS for Windows Version 12.0. SPSS, Inc., Chicago, Illinois.

Sutherland, W. J. 1996. Ecological census techniques. Cambridge University Press, Cambridge.

Velleman, P. F. 1997. DataDesk Version 6.0: Statistics Guide. Data Description. Ithaka, NY.

Vladykov, V. and E. Kott. 1980. Description and key to metamorphosel spawners and ammocoetes of Petromyzonidae found in the Great Lakes region. Canadian Journal of Fisheries and Aquatic Sciences. 376: 1616-1625.

Wiley, M. J., P. W. Seelbach, K. Wehrly, and J. Martin. 2003. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. Chapter 12 in: T. P. Simon, Editor. Biological response signatures: indicator patterns using aquatic communities. CRC Press, Boca Raton, Florida.

Chapter 5 : Biases arising from benthic macroinvertebrate methods in rapid bioassessment programs (RBPs) for wadeable streams: Michigan versus Korean assessment protocols

Abstract

Analyses of pooled datasets from different sources without appropriate validation of statistical comparability may result in critical bias and erroneous conclusions. In this study I compared benthic macroinvertebrate data and metrics collected using RBPs representative of those typically employed in North American and East Asian stream assessment studies. Specifically I explored similarities and differences between the Michigan (USA) Department of Environmental Quality Procedure 51 (MDEQP51) and the Korean Nationwide Aquatic Ecological Monitoring Program (KNAEMP). Paired sampling with these two methods was conducted at 29 stream sites in the Geum River Watershed, South Korea and their sampling performance was statistically compared with common benthic macroinvertebrate metrics. The two sampling methods target different habitats and use different sample sizes so that significant differences in taxa richness and individual numbers were frequently observed. In general, MDEQP51 method collected significantly lower taxa counts and individual numbers than KNAEMP method. However, indicator metrics showing relative percentages of EPT taxa and EPT individuals showed no significant difference, except for the species level metric. Numerical rescaling of macroinvertebrate data reduced the differences due to RBPs and led to comparable biological responses to land use stressor gradients.

Introduction

Protection of water resources has become a major societal priority during the last several decades (Barbour et al. 1999, Dodds 2006, Riseng et al. 2011, Allan et al. 2013). In response, numerous water-quality monitoring and assessment programs have been developed or adopted by both public and private organizations, states, and countries. Depending on objectives, operating constraints (budget and time), and crew familiarity with sampling methods (Diamond et al. 1996, Barbour et al. 1999, Carter and Resh 2001, Bonar and Hubert 2002), these programs employ differing sampling designs and often focus on different but restricted sets of the numerous biological indicators that are available. Water quality monitoring information derived from RBPs aims to answer questions about the condition of specific sites, changes over time, diagnosis of causes, and evaluation of remediation or prevention policy (ITFM 1995, Merritt and Cummins 1996, Barbour et al. 1999, Cao et al. 2005). Thus, the endpoints of all RBPs are to maintain healthy ecological conditions and to improve the quality of human life by promoting sustainable use of natural resources (ITFM 1995, MEA 2005, Furse et al. 2006, US EPA 2007).

Sampling-based methodologies used to monitor environmental quality vary in respect to efficiency, precision, and bias. Differences in methodologies can lead to biases that influence results and thus data comparability, and will have implications for meta-analysis of monitoring data obtained (ITFM 1995, Wiley et al. 2003, Furse et al. 2006). Monitoring methodologies are typically designed to meet regional needs and constraints (Carter and Resh 2001, Wiley et al. 2003, Cao et al. 2005), However, recently there have been number of comparative, and/or large-scale (including international) assessments which can require pooling of RBP data from different sources (ITFM 1995, MEA 2005, Furse et al. 2006, US EPA 2007, Riseng et al. 2010, Allan et al. 2013). In this context methodological variations in efficiency and bias of course become a concern (ITFM 1995, Barbour et al. 1999, Houston et al. 2002, Wiley et al. 2003).

Earlier studies have examined benthic macroinvertebrate-based rapid bioassessment methods to evaluate capture accuracy, efficiency and bias of sampling gear (Freeman et al. 1984, Stark 1993), degrees of field sampling effort (Stark 1993, Barbour and Gerritsen 1996, King and Richardson 2002, Park 2007), types of habitats sampled (Stark 1993, Parsons and Norris 1996, Rinella and Feminella 2005, Gerth and Herlihy 2006), levels of taxonomic resolution for identification (King and Richardson 2002, Park 2007), selection of indicator metrics (Klemm et al. 2003, Dahl and Johnson 2004, Blocksom et al. 2002), and effects of seasonal variation (Merritt and Cummins 1996, Sporka et al. 2006). These studies all examined the impacts of methodology within specific monitoring programs or regions. Studies have less frequently examined methodological issues related to sampling or assessment data from different monitoring and assessment programs (Cao et al. 2005, Clarke et al. 2006, Clarke and Hering 2006, Friberg et al. 2006). For example, the EU Water Framework Directive (EU-WFD) examined existing rapid bioassessment programs (RBPs) in European countries and compared them to a newly developed standard protocol developed to evaluate aquatic ecosystem health in European streams and rivers (Furse et al. 2006). The Environmental Monitoring and Assessment Program (EMAP) of the US Environmental Protection Agency (US EPA) also developed a standard RBP to assess status and trends of national ecological resources in the United States (US EPA 2007). Development and application of a single RBP across many states or nations can be inherently problematic because regional differences in geomorphological and biological heterogeneity may result in geographical biases even if a single standardized method is used. The

Intergovernmental Task Force on Monitoring Water Quality (ITFM) emphasized the importance of understanding data comparability from various RBPs (ITFM 1995).

Benthic macroinvertebrate sampling methods used by resource agencies in Michigan, USA and South Korea were compared here as a case study in assessment data comparability and integration. Michigan Department of Environmental Quality Procedure 51 (abbreviated below as MDEQP51) sampling is a fixed-count approach (e.g. 100 individuals) that is adopted from a standard operating procedure within the USEPA's RBPs (MDEQ 1997, Barbour et al. 1999). MDEQP51 samples characterize the structure of benthic macroinvertebrate communities in terms of relative abundances of each taxon rather than absolute density (Moulton et al. 2002). Due to the cost-saving benefits of fixed-count processing, RBPs using this approach have been preferred by many managers and biologists (Barbour and Gerritsen 1996, Walsh 1997, Doberstein et al. 2000, King and Richardson 2002). In contrast to Michigan's fixed-count approach (MDEQP51), the Korean Nationwide Aquatic Ecological Monitoring Program (KNAEMP below) sampling uses a fixed-area subsampling method in stream health assessment (King and Richardson 2002, NIER 2009) and is adapted from European RBPs based on the concept of saprobity (Sladecek 1973, Hellawell 1986, Furse et al. 2006). Saprobity refers to the physiological and biochemical characteristics of an organism that permit it to live inwater with some amount of organic matter (i.e., some degree of pollution) (Sladecek 1973). KNAEMP procedures estimate the absolute density and composition of benthic macroinvertebrate samples (NIER 2009, Jun et al. 2011). Walsh (1997) and King and Richardson (2002) reported that fixed-area subsampling was less efficient than fixed counts due to high variability of taxa and individual numbers collected, but improved both precision and accuracy of sampling data. These two different macroinvertebrate sampling approaches are currently used in many different regions and countries.

This study investigated the feasibility of comparing benthic macroinvertebrate data from these two very different stream RBPs (MDEQP51 and KNAEMP). My main goal was to study how sampling methods of two RBPs affected sampling performance, and to analyze what inherent potential biases in assessment metrics could be from the two RBPs. Finally, I examined whether or not the datasets from two different RBPs were able to produce the same stressor-responses relationships. My study provides an objective comparison of benthic macroinvertebrate data sampling with these two very different RBPs, and contributes to our knowledge about effective methodologies for ecological assessment of streams and rivers.

Materials and methods

Study sites and periods

Benthic macroinvertebrate sampling was conducted at 29 stream sites in April 2012 at the Geum River Watershed, South Korea (Figure 5.1) with two different rapid bioassessment methods. The methods I employed are commonly used in stream RBPs of South Korea (KNAEMP, NIER 2009) and Michigan, USA (MDEQP51, MDEQ 1997). The Geum River Watershed, located in the Midwestern part of South Korea, is a 244 miles (393 kilometer) long and drains 6,771 square miles (17,537 square kilometers; WAMIS 2013). This study watershed included various land use patterns and stream characteristics directly and indirectly affecting the variety of benthic macroinvertebrate habitat (Barbour et al. 1999, Wang et al. 2001, MDNR 2002, Wang et al. 2003, NIER 2009, Riseng et al. 2011). Sampling sites (n= 29) were a subset of the current KNAEMP macroinvertebrate sampling sites (total 130 sites for the Geum River Watershed, NIER 2009), which were previously designated for nationwide stream monitoring studies. One KNAEMP site was re-located nearby due to channel construction on the day of sampling.

Descriptive statistics for stream habitat variables and environmental factors of each sampling site (Table 5.1) included: catchment area, stream order, proportions of land use types, wetted stream width, average flow depth and velocity, water temperature, and proportions of substrate types. Stream order was calculated by Strahler's method (Strahler 1957) using a map scaled at 1:50,000 (NIER 2009). Water temperature and basic water chemistry (dissolved oxygen, pH, conductivity, and turbidity) was measured using a multi-parameter water quality sensor (YSI Environmental Monitoring System 660, Yellow Springs, OH, USA). Catchment areas and proportions of land use types for each site were summarized at the catchment scale from a digital map of watershed and 2004 land-cover/land-use using ArcGIS 9.1 (ESRI 2009). The digital map of watershed and land cover/land use was obtained from the WAter Management Information System (WAMIS 2013) of NIER, South Korea. Proportions of substrate types were estimated by visual examination of the coverage of each particle size class based on the modified Wentworth scales (Cummins 1962). Substrate types were classified into five classes: boulders (>256 mm in diameter), cobbles (64-256 mm in diameter), pebbles and gravels (2-64 mm in diameter), sand (0.06-2 mm in diameter), and clay and silt (<0.06 mm in diameter).

Benthic macroinvertebrate sampling

Each site was sampled with two different macroinvertebrate collection methods from Korean and Michigan RBPs to test the method-dependent differences in estimates of taxa richness, abundance, and other assessment metric values. The sampling was conducted simultaneously by two different crews working within the same stream reach. KNAEMP sampling, as prescribed in their protocol (NIER 2009), was conducted only in riffle areas, while MDEQP51 sampling (MDEQ 1997) occurred throughout the reach starting downstream and proceeding to the upstream limit. However, MDEQP51 crews did not collect samples in riffle areas until KNAEMP sampling was completed. Sampling reach distance for each site was generally 20 times the stream width of each site, but 50 meters was the minimum for small streams. Actual sampling reach length ranged from 50 meters to 100 meters and the actual sampling distance was determined to include multiple examples of the most common habitat units (riffles, runs, and pools).

Michigan department of environmental quality procedure 51 (MDEQP51) sampling

MDEQP51 samples characterize the structure of benthic macroinvertebrate communities in terms of relative abundances of each taxon rather than absolute density (Moulton et al. 2002). Sampling for this study was conducted in accordance with the Michigan Department of Environmental Quality Procedure 51 (MDEQ 1997). Sampling of benthic macroinvertebrate assemblages was performed using D-frame dip nets (250 μm mesh size) for 30 minutes at each site by one person. Kicking, dipping, and sweeping were used for general sampling with the dip net, and hand-picking was used for areas with boulders, debris, and logs. Samples from all habitats were combined in a bucket and then 100 organisms were randomly selected from the composite sample for further analysis (Merritt and Cummins 1996, MDEQ 1997, Riseng et al. 2006). The 100 selected organisms were preserved in 70 % ethanol and returned to the laboratory for identification and enumeration (Merritt and Cummins 1996). At a site, more than 200 organisms were collected due to picking error, so metric values were recalculated to fit a 100-organism scale.

Korean nationwide aquatic ecological monitoring program (KNAEMP) sampling

In contrast to MDEQP51 samples, KNAEMP samples estimate the absolute density of benthic macroinvertebrate communities. The density of benthic macroinvertebrates was calculated as individuals per square meter (NIER 2009, Jun et al. 2011). KNAEMP samples were quantitatively collected at riffle habitats using a Surber sampler (30 cm × 30 cm, 1 mm mesh size). Three samples at each site were taken from randomly selected riffles in a designed stream reach and placed into a 500 ml plastic bottle after removing large substrates and debris. Then 70 % ethanol was added to preserve samples for further identification and enumeration. KNAEMP sampling was conducted in accordance with the guidelines of the “National biological surveys for stream ecosystem health” in Korea (NIER 2009).

Identification of benthic macroinvertebrates

All organisms were separated from detritus and small substrate particles and sorted by each order. Then, all individuals were identified to the lowest taxonomic resolution level, typically species or genus. The identification of non-insects (Platyhelminthes, Nematomorpha, Mollusca, Annelida, and Crustacea) was done according to Kwon et al. (2001) and Smith (2001), while aquatic insects were based on Yoon (1995), Merritt and Cummins (1996), and Won et al. (2005). All identified individuals counted at the lowest taxonomic level. If lab identification was unsure or impossible, a few samples of each taxon from one site were preserved with 70% alcohol and labeled with date, site, and taxa name for later identification by taxonomic experts.

Benthic macroinvertebrate metrics and data analysis

All samples were evaluated for fourteen indicator metrics: total number of orders (*nOrTa*), total number of families (*nFaTa*), total number of genera taxa (*nGeTa*), total number of species (*nSpTa*), total number of individuals (*nTotIn*), total number of EPT families (Ephemeroptera, Plecoptera, and Trichoptera) (*nEPTFaTa*), total number of EPT genera (*nEPTGeTa*), total number of EPT species (*nEPTSpTa*), percentage of EPT families (*pEPTFaTa*), percentage of EPT genera (*pEPTGeTa*), percentage of EPT species (*pEPTSpTa*), percentage of EPT individuals (*pEPTIn*), Korean Saprobic Index (*KSI*) score, and Macroinvertebrate Biotic Index (*MBI*) score (Table 5.2). The scores of KSI and MBI have a negative relationship with ecological integrity, which means higher values indicate poor ecological status and degradation in water quality. Korean saprobic index is a modified benthic macroinvertebrate index of biological integrity from the saprobic valency concept (Zelinka and Marvan 1961) and is currently used for the ecological assessment of Korean streams and rivers (NIER 2009). Saprobic values and weighting factors were summarized for 100 major benthic macroinvertebrate indicator groups. KSI score ranges from 0 (excellent condition) to 5 (poor condition) and the KSI score of each site was calculated by averaging sum of Saprobic value and weighting factor of each taxon collected (Won et al. 2006 and NIER 2009). Macroinvertebrate Biotic Index is taken from Hilsenhoff or EPA established biotic index values (Hilsenhoff 1987 and USEPA 2006). A tolerance value for each taxon ranged from 0 to 10 and the average MBI score of each site was calculated by averaging the sum of a published tolerance value for each taxon collected (Riseng et al. 2006). These indicator metrics were chosen to describe representative measures of richness and composition for a variety of macroinvertebrate assemblages (MDEQ 1997, Barbour et al. 1999, NIER 2009).

Benthic Macroinvertebrate Sampling Method Efficiency (SME) was estimated for seven key metrics, counting taxa richness (Table 5.2). Also, another fourteen metrics were analyzed in order to observe the difference in taxa richness and relative abundance amongspecific major groups: total number of class taxa (*nClTa*), total number of EPT individuals (*nEPTIn*), total number of Ephemeroptera individuals (*nEphIn*), total number of Plecoptera individuals (*nPleIn*), total number of Trichoptera individuals (*nTriIn*), total number of Ephemeroptera families (*nEphFaTa*), total number of Ephemeroptera genera (*nEphGeTa*), total number of Ephemeroptera species (*nEphSpTa*), total number of Plecoptera families (*nPleFaTa*), total number of Plecoptera genera (*nPleGe*), total number of Plecoptera species (*nPleSpTa*), total number of Trichoptera families (*nTriFaTa*), total number of Trichoptera genera (*nTriGeTa*), and total number of Trichoptera species (*nTriSpTa*). SME here can be defined as follows,

$$SME = 100 \times (SM_{ED} / SR_{CB}) \quad (1)$$

where SM_{ED} indicates taxa richness and relative abundance of each data (SM_{ED}) set for each sampling method and SM_{CB} indicates taxa richness and relative abundance of Combined data set (SM_{CB}) from both KNAEMP and MDEQP51 data. CB data set was used as a maximum taxa richness and relative abundance of each sampling site because we could not measure total taxa list and abundance for each site.

Standardization was performed for each benthic macroinvertebrate metric separately in order to examine whether differences and biases from sampling methods can be corrected. The absolute value of z represents the distance between the raw score and the metric mean in units of standard deviation. The standard score of a raw score x is

$$z = (x - \mu)/\sigma \quad (2)$$

where μ is the mean of each metric for each sampling method and σ is the standard deviation of each metric for each sampling method.

Statistical analysis

Benthic macroinvertebrate data sets were summarized into three categories to compare the sampling performance of sampling methodologies: KNAEMP, MDEQP51, and COMB (combined data set of KNAEMP and MDEQP51). Statistical summaries (mean, median, standard deviation, minimum, and maximum) and scatter plots were conducted using Datadesk 6.0 (Velleman 1997). Paired samples t-test, ANCOVA test, Pearson correlation, and Standardization of data sets were performed using SPSS 12.0 (SPSS, Inc. 2003).

Results

Characteristics of stream sites and benthic macroinvertebrate habitats

Environmental and watershed characteristics (Table 5.1) showed large variation across sampling sites. Catchment size ranged from 16 km^2 to 8,712 km^2 with average of 1,382 km^2, representing sites from the Geum headwaters sites to large mainstem river reaches. The Geum River is a 6th order river system and stream order for this study ranged from 2nd order to 6th order with average of 4th order. Mean proportions of urban, agricultural, and forest land use were 0.15, 0.27, and 0.50 and maximum proportions were 0.78, 0.81, and 0.87, respectively. Wetted stream widths ranged 0.5 meter to 250.0 meter with average of 56.2 meter. Average flow depths ranged from 0.13 m to 0.54 m with average of 0.33 m and average velocity ranged from zero to 1.24 m/sec with an average of 0.60 m/sec. Water temperature also showed large variation, ranging from 5.2 degree Celsius to 14.0 degree Celsius with an average of 10.4 degree Celsius on the sampling date. Mean percentages of substrates for clay and silt, sand, gravels, cobbles, and boulders were 13.3%, 22.4%, 22.0%, 23.1%, and 19.2%, respectively.

Comparison of sampled taxa richness, abundance, and occurrence

Sampled benthic macroinvertebrates (combined data set; COMB) collected by both MDEQP51 and KNAEMP included 20 orders, 64 families, 107 genera, 138 species, and 17,878 individuals (Table 5.3). Ephemeroptera had the highest species richness (41 species) followed by Trichoptera (24 species), Diptera (18 species), and Plecoptera (10 species). For total individual numbers Diptera had the highest individuals (8,422) followed by Ephemeroptera (3,826), Trichoptera (3,751), Archioligocheata (845), Isopoda (312), Coleoptera (296), and Plecoptera (126) (Table 5.3 and Figure 5.2).

In general, the KNAEMP method captured more species taxa and individuals than the MDEQP51 method, whereas total species taxa richness of Odonata and Hemiptera for MDEQP51 was relatively higher than those of the KNAEMP (Table 5.3). Also, total individual numbers of Hemiptera in MDEQP51 method were higher than in the KNAEMP method. In terms of percentages of combined data, the KNAEMP method sampled 95.0% of total orders (19 orders), 90.6% of total families (58 families), 78.5% of total genera (84 genera), 83.3% of total species (115 species), and 82.5% of total individuals (14, 734 individuals). In contrast, the MDEQP51 method captured 85.0% of total orders (17 orders), 71.9% of total families (46 families), 61.7% of total genera (66 genera), 59.4% of total species (82 species), and only 17.6% of total individuals (3,144 individuals). Interestingly, Trichoptera in KNAEMP method showed much higher total taxa richness (100.0% of total families (11 families), 87.5% of total genera (14 genera), 91.7% of total species (22 species)) and total individual numbers (92.3% of total individuals (3,461 individuals) than those of Trichoptera in MDEQP51 method (45.5% (5), 50.0% (8), 45.8% (11), and 7.7% (290), respectively). However, Odonata in KNAEMP method showed relatively lower total taxa richness (50% of total families (2 family), 33.3% of total genera (3 genera), and 33.3% of total species (3 species)) than total taxa richness (100% (4), 88.9% (8), and 88.9% (8)) of Odonata in MDEQP51 method.

Site occurrence frequencies and relative abundances of benthic macroinvertebrate groups were well described by both sampling methods (Figure 5.2). In terms of relative abundances the KNAEMP method showed good performance for taxa groups in riffle habitats, while the MDEQP51 method generally captured various taxa groups (e.g., Odonata, Isopoda, and

Hemiptera) in diverse habitats (runs and pools) in addition to riffle areas. In terms of site occurrence frequencies of orders, almost no differences were observed between two sampling methods. The KNAEMP method indicated that Diptera had the highest site occurrence (29 sites) followed by Archioligocheata (25 sites), Ephemeroptera (25 sites), Trichoptera (24 sites), Tricladida (14 sites). In contrast the MDEQP51 method showed that Diptera had the highest occurrence (29 sites) followed by Ephemeroptera (24 sites), Trichopetera (23 sites), Archioligocheata (23 sites), and Plecoptera (10 sites).

Benthic macroinvertebrate metric scores based on taxa counts and individual numbers (Table 5.4) were relatively higher in KNAEMP samples than MDEQP51 samples for all taxonomic levels. For example, the mean total taxa numbers of order, family, genus, and species in the KNAEMP samples were 6.8, 13.1, 15.9, and 19.0, while the means in the MDEQP51 samples were 5.7, 9.7, 11.5, and 14.4, respectively. These patterns were continuously observed in the indicator metrics based on EPT taxa counts and individual numbers at all taxonomic resolution. However, KSI and MBI (overall assessment metrics) showed the opposite trend. KSI and MBI values from the MDEQP51 sampling method produced relatively higher values than the KNAEMP sampling method. Mean KSI and MBI in MDEQP51 samples were 2.12 (0.42 to 4.56) and 6.20 (4.13 to 9.92), respectively; whereas the means from KNAEMP samples were 1.73 (0.28 to 5.00) and 5.88 (4.01 to 9.44), respectively.

Comparison of macroinvertebrate sampling method efficiency (SME)

For benthic macroinvertebrate metrics based on taxa counts, KNAEMP method showed excellent SMEs for all metrics (Figure 5.3A). The mean SMEs of the KNAEMP method ranged from 78.3% (*nEPTSpTa*) to 90.4% (*nClTa*). However, SMEs in the MDEQP51 method were significantly lower ranging from 54.1% (*nEPTSpTa*) to 74.3% (*nOrTa*). A similar pattern was found for indicator metrics based on individual numbers (Figure 5.3B). Mean SMEs of the KNAEMP method ranged from 54.6% (*nPleIn*) to 87.4% (*nTriIn*), whereas mean SMEs of the MDEQP51 method were significantly lower (except nEphln) than those of KNAEMP method, ranging from 12.6% (*nTriIn*) to 45.4% (*nPleIn*).

There were also interesting differences in relative performance of the methods between Ephemerotera, Plecoptera, and Trichoptera orders (Figure 5.4). For all taxonomic levels, mean MSMEs of Trichoptera with taxa counts were significantly higher (89.8% to 92.8%) in KNAEMP method than those (39.2% to 44.8%) of MDEQP51 method, as this pattern was observed in the analysis of *nTriIn* (Figure 5.3B). However, no significant difference by sampling method was observed in mean MSMEs for Ephemeroptera and Plecoptera taxa counts at all taxonomic levels (Figure 5.4).

Metric comparability

Pearson correlation analysis indicated that most of metrics between sampling methods were significantly correlated ($p<0.01$) ranging from 0.563 (NIN) to 0.937 (*nEPTSpTa*) (Table 5.5). Benthic macroinvertebrate metrics measuring various aspects of taxa richness (seven metrics) and individual numbers (one metric) were significantly different between KNAEMP and MDEQP51 methods (Table 5.6). Regression slopes of KNAEMP and MDEQP51 method data ranged from 1.1416 (*nOrTa*) to 1.3763 (*nEPTGeTa*), indicating that the KNAEMP method had consistently higher values than the MDEQP51 method for these metrics (Figure 5.5). Regression slope for NIN was not calculated because MDEQP51 method only sampled 100 macroinverbrate organisms, which was inappropriate to compare.

Benthic macroinvertebrate metrics based on relative percentages of EPT taxa counts (three metrics) and individual numbers (one metric) showed somewhat different results from taxa-count based metrics (Table 5.5, Table 5.6, and Figure 5.6). The metrics for *pEPTFaTa*, *pEPTGeTa*, and *pEPTIn* were not significantly differ ($p>0.05$) with sampling methods (Table 5.6). The exception was observed in *pEPTSpTa*, which showed a strongly significant difference ($p<0.05$). The Pearson correlation tests showed strong correlation between KNAEMP and MDEQP51 samples for all four of these metrics (Table 5.5), with correlation coefficients ranged from 0.667 (*pEPTFaTa*) to 0.893 (*pEPTIn*). Regression slopes (Figure 5.6) for these indicator metrics ranged from 0.9710 (*pEPTGeTa*) to 1.0432 (*pEPTSpTa*), not significantly different from 1.

Overall stream health metrics (*KSI* and *MBI*) varied in sensitivity to sampling method. *MBI* scores had a strongly significant difference ($p=0.015$) between sampling methods, whereas *KSI* scores were not significantly different ($p=0.075$) although the p value was marginal (Table 5.6). Pearson correlation indicated that *MBI* ($r=0.899$) and *KSI* ($r=0.726$) values were well correlated between sampling methods (Table 5.5). Regression slopes (Figure 5.6) of KNAEMP and MDEQP51 method data were 0.9366 and 0.8148 for *MBI* and *KSI*, respectively, indicating that very closed to 1 and regression slopes were not significantly different (ANCOVA test; $p=0.362$ for *MBI* and $p=0.333$ for *KSI*) between methods.

In the Geum River basin, site bioassessment metric values and ecological health had significantly negative correlaion with both agricultural and urban land use (Figure 5.7 and Figure 5.8). Sampling methods influenced the sensitivity of diversity-related metrics (i.e. those based on taxa counts) and as a result slopes of the regression coefficients (ANCOVA) of several typical metrics differed significantly by sampling methodology (Table 5.7). Particularly, *nFaTa*, *nGeTa*, and *nTotIn* metrics were significantly influenced by sampling methods, while other metrics were marginally significant. This nicely illustrated the impact of methodological biases on comparative assessments of sensitivity to landscape stressors (Figure 5.7 and Figure 5.8). In this case the effect varied between metrics, presumably reflecting different sensitivities to both sample size and spatial extent of the rapid bioassessment protocol. Calibration-based corrections or statistical normalization (as discussed below) would be required to validly pool or compare metrics generated by these different sampling methods (Figure 5.7 and Figure 5.8).

Discussion

Benthic macroinvertebrate data collected using typical North American (MDEQ 1997) and East Asian (NIER 2009) rapid assessment methodologies were compared in order to examine the feasibility of direct metric value comparison. This study particularly focused on the types of potential biases that might be associated with the two programs. Significant differences in sampling bias would complicate direct comparisons of assessment results from these regions, even if the same specific assessment metrics are used. Undoubtedly all sampling gear and all sampling protocols have inherent biases with respect to size of individuals captured, efficiency in various habitats, and taxonomic representation (Barbour et al. 1999, Wiley et al. 2003, Furse et al. 2006). As long as a single sampling protocol is maintained, it is reasonable to assume that whatever those biases are, measurements using a given sampling protocol are comparable in a general sense (but see Larimore 1961, Wiley et al. 2003). But when data collected using different methods are compared, biases can lead to quite different representations of the same sampled community (Barbour et al. 1999, Carter and Resh 2001, Wiley et al. 2003, Cao et al. 2005, ITFM 2005, Clarke et al. 2006, Furse et al. 2006).

I found that the two sampling methodologies I investigated resulted in significant systematic differences in taxa and individual specimen counts and therefore in the metric values that were based on those counts. The KNAEMP method consistently recovered more taxa than MDEQP51 method across all indicator metrics (Table 5.4). This is not surprising since the MDEQP51 protocol typically ends with the random selection of only 100 individuals for identification. In contrast the KNAEMP method uses actual counts from surber samples and therefore reflects prevailing benthic densities which typically range in streams and rivers up to several thousand per square meter. Since sample diversity is related to sample size (Merritt and

Cummins 1996, Walsh 1997, Barbour et al. 1999, Carter and Resh 2001, King and Richardson 2002), sampling protocol had a significant effect on count-based metrics; both metrics based on individual counts and those based on taxa counts (e.g. number of EPT taxa and total number of genera). However, percentage based metrics were unaffected (Table 5.5). This influence of sample size is conflated in the protocols with differences related to spatial scale of the samples. Since the two sampling methods targeted somewhat different habitats, efficiency of sampling certain taxonomic groups associated with those habitats also varied with sampling protocol. For example the KNAEMP method sampled significantly more Trichoptera taxa, many of which typically reside in high densities only in riffle habitats (e.g. the hydrosychoidea, brachycentrids, glossosmatids, goerids), whereas the MDEQP51 method systematically captured more taxa from non-riffle habitats like pools and edges (e.g., Odonata and Hemiptera). Overall these results imply that common indicator metrics based on counts of taxa or individuals should not be compared directly if they come from sampling methods with differing sampling efficiencies and biases (Carter and Resh 2001, Cao et al. 2005, ITFM 2005, Friberg et al. 2006, Furse et al. 2006). Park (2007) reported the similar results in a study of the effects of sampling effort on stream assessment metrics. Such count-based metrics are likely to require some form of calibration or normalization process (Wiley et al. 2003, Park 2007) before comparison or integration of data sets.

The two bioassessment sampling protocols examined in this study also have adopted different taxonomic levels for benthic macroinvertebrate identification and stream health assessment (MDEQ 1997, NIER 2009). MDEQP51 method uses family level identification, whereas KNAEMP method applies species level identification. Level of taxon identification can obviously affect lab analysis time and project budget (Merritt and Cummins 1996, Barbour et al.

1999, King and Richardson 2002, Park 2007). In this study species-level identification for benthic macroinvertebrates required several times more effort than either family- or genus-level identification. From the perspective of program design, fast and easy identification is a significant benefit which allows much useful work to be performed by public users with brief training or guide books (Barbour et al. 1999, ITFM 2005, Furse et al. 2006, Park 2007, US EPA 2007). Furthermore, species-level identification clearly has higher time and labor costs; and also requires more expertise and more equipment used for identification of specific benthic macroinvertebrate taxa. On the other hand, metrics using the KNAEMP methodology, including KSI generally had a wider range in values which potentially yields a more sensitive metric. In terms of simple linear sensitivity to both urban and agricultural development in the upstream catchment, this was indeed the case: KNAEMP metrics were generally more sensitive than MDEQP51-based metrics (see Figure 5.7, Figure 5.8).

Indicator metrics (four metrics) based on relative percentages of taxa counts and individuals did not differ significantly between KNAEMP and MDEQP51 samples (Table 5.4, Table 5.5, Table 5.6), implying that EPT indicator metrics based on relative percentages were not influenced by sampling methods or sample sizes. Thus, EPT indicator metrics based on relative percentages of taxa counts may be preferable to taxa-count based metrics when combining or comparing assessment metrics from different sources. However, the percentage of EPT taxa metric based on the species level still showed statistically significant differences indicating that the larger sample sizes (n) from the KNAEMP method still influenced relative percentages of EPT at the species level.

Methodological difference in sampling target habitats between two sampling methods was also reflected in EPT indicator metrics (seven metrics) and stream assessment metrics (KSI and MBI). KNAEMP method had slightly higher means for all EPT metrics than MDEQP51 method. Particularly, metrics based on taxa counts were significantly different at every taxonomic resolution ($p<0.01$). Similarly, mean values of KSI and MBI were lower in KNAEMP method than MDEQP51 method. In general, EPT taxa are intolerant to environmental changes so that these taxa have lower tolerance scores. More EPT taxa counts in KNAEMP method than MDEQP51 method were expected given the larger sample sizes. However, relative percentages of EPT taxa counts and individuals could be similar if they were sampled in the same habitat areas regardless of sampling effort or efficiency. Although metrics based on relative percentages of EPT taxa and individuals were not significantly different ($p>0.05$) except for *pEPTSpTa* ($p=0.036$), the slightly higher means in KNAEMP method likely reflects a methodological focus on their preferred habitats. It is well known that riffle habitats provide more oxygen and turbulence and are generally highly suitable for EPT and sensitive taxa groups (Merritt and Cummins 1996, Diana 2004, Allan and Castillo 2007).

Assessing impacts with biased indicators

All of the assessment metrics, regardless of the sampling protocol employed, were statistically correlated with both agricultural and urban land use and so reflect environmental gradients. However, biases related to sampling methodology led to a statistically different stressor-response relationship in many cases (see Figure 5.7, Figure 5.8). This is a critically important issue for my dissertation, since my goal is to compare stressor-response relationships from two different regions historically based on different sampling protocols and metrics. In rapid bioassessment studies potential biases of individual metrics are often addressed through the use of multi-metrics (e.g. Karr 1981, Karr et al. 1986, Ohio EPA 1987, Barbour et al. 1999) reflecting a “measurement model” (Blalock 1970) approach to overcoming the limitations of individual metrics. However when bias is methodological, multiple metrics calculated from the data share the bias and pooling indicators simply pool the biases they carry (e.g. Table 5.6). Gear bias in fisheries studies is sometimes addressed through direct calibration and correction factors (see chapter 4). Data standardization (normalization) is another useful statistical approach when comparing data with different ranges and or known biases (Wiley et al. 2003, Baker et al. 2005, Riseng et al. 2006, Park 2007, Riseng et al. 2010, Launois et al. 2011). Standardizing my Geum river sample sets independently (i.e. for KNAEMP and MDEQP51 samples separately) successfully removed the difference between sample types in stressor-response relationships (Table 5.7, Figure 5.7, Figure 5.8); the different sampling methods provided statistically identical stressor-response relationships for the same set of sites as they should. Thus, standardization is likely a necessary step before comparison of macroinvertebrate metrics when the underlying sampling methodologies are quite different.

Summary and conclusion

Benthic macroinverbrates are often used in stream monitoring and assessment studies due to the convenience and low cost of field sampling (Merritt and Cummins 1996, Barbour et al. 1999, Furse et al. 2006, US EPA 2006, Allan and Castillo 2007). Various sampling gears and methods for benthic macroinvertebrates are available and have been selectively used in many states and countries, depending on project purposes and operational constraints (ITFM 1995, Merritt and Cummins 1996, Barbour et al. 1999, Cao et al. 2005). However, the differences and biases inherent in sample data from common RBPs have seldom been addressed (Cao et al. 2005, Clarke et al. 2006, Clarke and Hering 2006, Friberg et al. 2006). In this study the MDEQP51 sampling approach frequently resulted in different metric scores than the KNAEMP procedure and exhibited larger variation in performance for certain benthic macroinvertebrate groups. The two RBPs examined resulted in significantly different land use stressor-response relationships for the same set of Geum River assessment sites. However, standardization of each data set by sampling type resolved this concern and successfully corrected the biological responses of each metric to land use stressors. MDEQP51 is more cost-effective than the KNAEMP, because the method reduces the taxonomic identification work. Although KNAEMP method requires more effort with higher labor costs, it provides much better estimates of density and diversity.

In this chapter I have examined the comparability of benthic macroinvertebrate data using different sampling methods as a step towards comparing land use stressor-response relationships in NA and East Asian regions. I conclude that macroinvertebrate data from different sampling methods are comparable after appropriate numerical calibration.

Table 5.1. Summary statistics of stream habitat characteristics and landuse patterns for benthic macroinvertebrate sampling sites. SD indicates standard deviation.

Variables (n= 29)	Mean	Median	SD	Min	Max
Catchment area (km^2)	1382	287	2501	16	8712
Stream order	4.2	4.0	1.3	2.0	6.0
Proportion of urban	0.15	0.06	0.21	0.00	0.78
Proportion of agriculture	0.27	0.22	0.19	0.00	0.81
Proportion of forest	0.50	0.57	0.25	0.09	0.87
Wetted stream width (m)	56.2	25.0	67.7	0.5	250.0
Average flow depth (m)	0.33	0.35	0.10	0.13	0.54
Average velocity (cms)	0.60	0.64	0.35	0.00	1.24
Water temperature (°C)	10.4	10.8	2.4	5.2	14.0
Percentage of clay and silt	13.3	10.0	16.7	0.0	90.0
Percentage of sand	22.4	10.0	19.7	5.0	70.0
Percentage of gravels	22.0	25.0	9.0	0.0	35.0
Percentage of cobbles	23.1	25.0	10.8	0.0	0.4
Percentage of boulders	19.2	20.0	12.6	0.0	45.0

Table 5.2. Macroinvertebrate indicator metrics and their abbreviations used for this study. * indicates that metrics were used for the evaluation of sampling gear efficiency.

	Metric list	**Abbreviation**
Metric 1	Total number of order taxa	*nOrTa**
Metric 2	Total number of family taxa	*nFaTa**
Metric 3	Total number of genus taxa	*nGeTa**
Metric 4	Total number of species taxa	*nSpTa**
Metric 5	Total number of individuals	*nTotIn*
Metric 6	Total number of EPT family taxa	*nEPTFaTa**
Metric 7	Total number of EPT genus taxa	*nEPTGeTa**
Metric 8	Total number of EPT species taxa	*nEPTSpTa**
Metric 9	Percentage of EPT family taxa	*pEPTFaTa*
Metric 10	Percentage of EPT genus taxa	*pEPTGeTa*
Metric 11	Percentage of EPT species taxa	*pEPTSpTa*
Metric 12	Percentage of EPT individuals	*pEPTIn*
Metric 13	Korean saprobic index	*KSI*
Metric 14	Macroinvertebrate Biotic Index	*MBI*

Table 5.3. Comparison of benthic macroinvertebrate data for specific orders. Data were summarized by MDEQP51, KAEMAP, and COMB (combined data of MDEQP51 and KAEMAP).

		nOrTa	***nFaTa***	***nGeTa***	***nSpTa***	***nTotIn***
MDEQP51	Overall	17	46	66	82	3144
	Order Ephemeroptera	1	9	18	27	747
	Order Plecoptera	1	4	5	5	49
	Order Trichoptera	1	5	8	11	290
	Order Odonata	1	4	8	8	14
	Order Coleoptera	1	3	3	3	23
	Order Hemiptera	1	3	3	3	4
	Order Megaloptera	1	1	2	2	7
	Order Diptera	1	7	9	11	1323
KNAEMP	Overall	19	58	84	115	14734
	Order Ephemeroptera	1	10	23	38	3079
	Order Plecoptera	1	5	8	10	77
	Order Trichoptera	1	11	14	22	3461
	Order Odonata	1	2	3	3	18
	Order Coleoptera	1	3	5	5	273
	Order Hemiptera	1	2	2	2	2
	Order Megaloptera	1	1	2	2	14
	Order Diptera	1	9	11	14	7099
COMB	Overall	20	64	107	138	17878
	Order Ephemeroptera	1	12	27	41	3826
	Order Plecoptera	1	4	8	10	126
	Order Trichoptera	1	11	16	24	3751
	Order Odonata	1	4	9	9	32
	Order Coleoptera	1	4	6	6	296
	Order Hemiptera	1	4	4	4	9
	Order Megaloptera	1	1	2	2	21
	Order Diptera	1	10	15	18	8422

Table 5.4. Summary statistics of all macroinvertebrate indicator metrics used for macroinvertebrate data comparability (n = 29).

Metrics	KNAEMP					MDEQP51					COMB				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
nOrTa	6.8	7.0	2.8	2.0	12.0	5.7	5.0	2.1	2.0	10.0	7.8	8.0	2.8	2.0	13.0
nFaTa	13.1	13.0	7.0	2.0	29.0	9.7	9.0	5.1	2.0	19.0	15.7	14.0	7.7	3.0	31.0
nGeTa	15.9	15.0	9.8	2.0	38.0	11.5	11.0	6.5	2.0	22.0	20.3	19.0	11.1	4.0	41.0
nSpTa	19.0	17.0	12.3	2.0	45.0	14.4	13.0	7.8	3.0	27.0	25.4	23.0	14.4	5.0	54.0
nTotIn	508.1	469.0	389.9	7.0	1,327.0	108.4	115.0	38.5	12.0	201.0	616.5	593.0	412.8	19.0	1,472.0
nEPTFaTa	6.8	7.0	4.5	0.0	16.0	4.9	4.0	3.5	0.0	11.0	7.7	7.0	4.9	0.0	17.0
nEPTGeTa	9.2	8.0	6.8	0.0	23.0	6.5	6.0	4.8	0.0	14.0	10.9	10.0	7.8	0.0	25.0
nEPTSpTa	11.7	10.0	9.0	0.0	27.0	8.0	6.0	6.7	0.0	19.0	14.4	11.0	11.2	0.0	34.0
pEPTFaTa	45.4	50.0	20.0	0.0	75.0	43.8	50.0	20.0	0.0	75.0	44.3	50.0	16.2	0.0	71.4
pEPTGeTa	49.5	55.6	21.6	0.0	81.3	48.1	55.0	22.2	0.0	81.8	47.0	52.4	18.4	0.0	77.8
pEPTSpTa	51.6	58.8	23.1	0.0	77.8	45.4	50.0	25.5	0.0	85.7	47.3	52.6	22.0	0.0	81.8
pEPTIn	32.2	25.6	28.4	0.0	85.4	30.6	16.7	28.7	0.0	91.1	32.1	22.4	28.1	0.0	86.9
KSI	1.73	1.09	1.57	0.28	5.00	2.12	2.18	1.42	0.42	4.56					
MBI	5.88	6.07	1.15	4.01	9.44	6.20	6.16	1.46	4.13	9.92					

Table 5.5. Two tailed Pearson correlation tests between KNAEMP and Michigan MDEQP51. Bold indicates significance at p≤0.05, and bold and italics indicate significance at p≤0.01.

	MDEQP51													
	nOrTa	*nFaTa*	*nGeTa*	*nSpTa*	*nTotIn*	*nEPTFaTa*	*nEPTGeTa*	*nEPTSpTa*	*pEPTFaTa*	*pEPTGeTa*	*pEPTSpTa*	*pEPTIn*	*KSI*	*MBI*
KNAEMP														
nOrTa	***.568***	***.647***	***.692***	***.683***	***.474***	***.698***	***.709***	***.697***	***.664***	***.645***	***.710***	***.553***	***-.617***	***-.491***
nFaTa	***.586***	***.803***	***.829***	***.837***	***.539***	***.875***	***.870***	***.874***	***.785***	***.767***	***.855***	***.773***	***-.745***	***-.659***
nGeTa	***.565***	***.814***	***.840***	***.855***	***.484***	***.905***	***.895***	***.901***	***.791***	***.767***	***.860***	***.797***	***-.778***	***-.700***
nSpTa	***.575***	***.825***	***.854***	***.865***	***.488***	***.908***	***.901***	***.908***	***.781***	***.758***	***.859***	***.810***	***-.792***	***-.719***
nTotIn	.335	**.437**	**.454**	**.444**	***.563***	***.500***	***.498***	***.517***	***.570***	***.581***	***.639***	***.490***	***-.558***	***-.467***
nEPTFaTa	***.544***	***.800***	***.829***	***.838***	***.532***	***.882***	***.886***	***.895***	***.808***	***.787***	***.888***	***.844***	***-.788***	***-.724***
nEPTGeTa	***.533***	***.815***	***.847***	***.863***	***.479***	***.918***	***.918***	***.930***	***.820***	***.796***	***.900***	***.868***	***-.823***	***-.763***
nEPTSpTa	***.541***	***.826***	***.862***	***.874***	***.479***	***.921***	***.922***	***.937***	***.806***	***.783***	***.897***	***.876***	***-.831***	***-.776***
pEPTFaTa	**.399**	***.514***	***.535***	***.524***	***.508***	***.581***	***.587***	***.585***	***.667***	***.649***	***.700***	***.572***	***-.603***	***-.521***
pEPTGeTa	**.428**	***.567***	***.596***	***.586***	***.523***	***.647***	***.657***	***.656***	***.728***	***.715***	***.775***	***.638***	***-.670***	***-.586***
pEPTSpTa	**.450**	***.601***	***.630***	***.620***	***.534***	***.679***	***.687***	***.691***	***.750***	***.739***	***.807***	***.665***	***-.697***	***-.620***
pEPTIn	***.498***	***.762***	***.809***	***.817***	***.495***	***.842***	***.864***	***.895***	***.748***	***.737***	***.867***	***.893***	***-.729***	***-.715***
KSI	-.196	**-.427**	**-.469**	***-.497***	-.197	***-.571***	***-.585***	***-.603***	***-.642***	***-.633***	***-.672***	***-.573***	***.726***	***.631***
MBI	***-.542***	***-.725***	***-.741***	***-.749***	-.238	***-.741***	***-.730***	***-.746***	***-.682***	***-.643***	***-.723***	***-.701***	***.818***	***.899***

Table 5.6. Summary statistics of paired samples T-tests (KNAEMP versus MDEQP51). CI and df indicate confidence interval and degree of freedom, respectively. *=0.05; **=0.001.

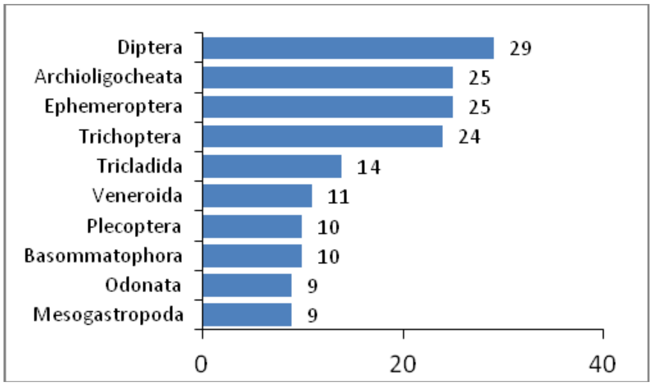
Metrics	Paired differences					t	df	Sig. (2-tailed)
	Mean	Standard deviation	Standard error mean	95% CI of the difference				
				Lower	Upper			
			KNAEMP versus MDEQP51					
nOrTa	1.103	2.335	0.434	0.215	1.992	2.545	28	0.017*
nFaTa	3.414	4.205	0.781	1.814	5.013	4.372	28	0.000**
nGeTa	4.379	5.628	1.045	2.239	6.520	4.190	28	0.000**
nSpTa	4.621	6.811	1.265	2.030	7.211	3.653	28	0.001**
nTotIn	399.655	369.665	68.645	259.042	540.268	5.822	28	0.000**
nEPTFaTa	1.862	2.150	0.399	1.044	2.680	4.664	28	0.000**
nEPTGeTa	2.690	3.060	0.568	1.526	3.854	4.733	28	0.000**
nEPTSpTa	3.621	3.610	0.670	2.248	4.994	5.402	28	0.000**
pEPTFaTa	1.582	16.244	3.016	-4.596	7.761	0.525	28	0.604
pEPTGeTa	1.315	16.554	3.074	-4.982	7.612	0.428	28	0.672
pEPTSpTa	6.253	15.284	2.838	0.439	12.067	2.203	28	0.036*=
pEPTIn	1.638	13.210	2.453	-3.387	6.663	0.668	28	0.510
KSI	-0.384	1.117	0.207	-0.809	0.041	-1.851	28	0.075
MBI	-0.317	0.659	0.122	-0.568	-0.067	-2.596	28	0.015*=

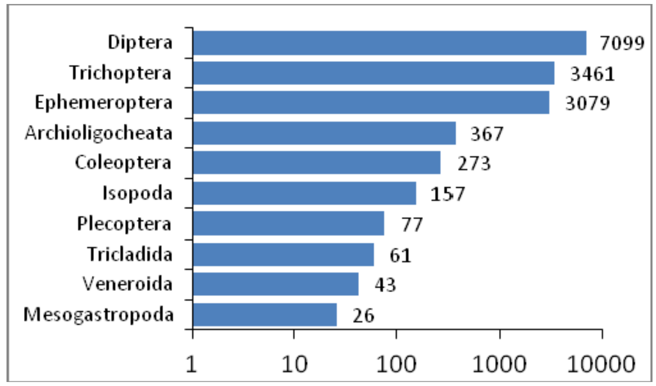
Table 5.7. P values from ANCOVA tests for benthic macroinvertebrate metrics. Each sampling method indicated KNAEMP and MDEQP51 and was used as fixed factor. Also, urban and agricultural land uses were used as covariates to observe biological response to land uses.

Metrics	Before standardization				After standardization			
	Method	Proportion of urban	Method	Proportion of agriculture	Method	Proportion of urban	Method	Proportion of agriculture
nOrTa	0.094	0.405	0.064	0.001	1.000	0.470	1.000	0.000
nFaTa	0.033	0.038	0.025	0.002	1.000	0.041	1.000	0.002
nGeTa	0.040	0.012	0.036	0.003	1.000	0.013	1.000	0.002
nSpTa	0.076	0.007	0.074	0.004	1.000	0.007	1.000	0.003
nTotIn	0.000	0.044	0.000	0.512	1.000	0.162	1.000	0.398
nEPTFaTa	0.066	0.006	0.067	0.007	1.000	0.005	1.000	0.006
nEPTGeTa	0.065	0.002	0.071	0.008	1.000	0.002	1.000	0.006
nEPTSpTa	0.062	0.001	0.073	0.009	1.000	0.001	1.000	0.008
pEPTFaTa	0.754	0.023	0.757	0.055	1.000	0.023	1.000	0.056
pEPTGeTa	0.812	0.014	0.816	0.051	1.000	0.014	1.000	0.053
pEPTSpTa	0.299	0.003	0.318	0.039	1.000	0.003	1.000	0.042
pEPTIn	0.814	0.002	0.820	0.018	1.000	0.002	1.000	0.018
KSI	0.291	0.001	0.326	0.108	1.000	0.001	1.000	0.106
MBI	0.315	0.001	0.335	0.007	1.000	0.001	1.000	0.006

Figure 5.1. Locations of benthic macroinvertebrate sampling sites (n= 29) in the Geum River Watershed, South Korea.

Macroinvertebrate data by KNAEMP

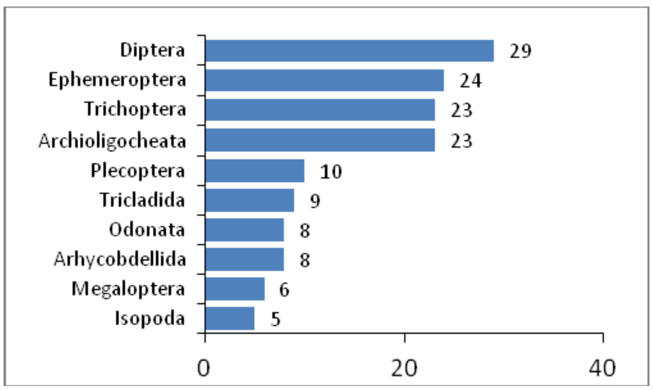


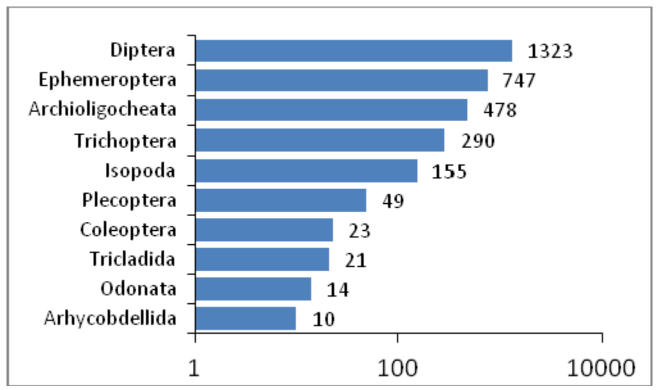


Site occurrence frequency

Relative abundance

Macroinvertebrate data by MDEQP51

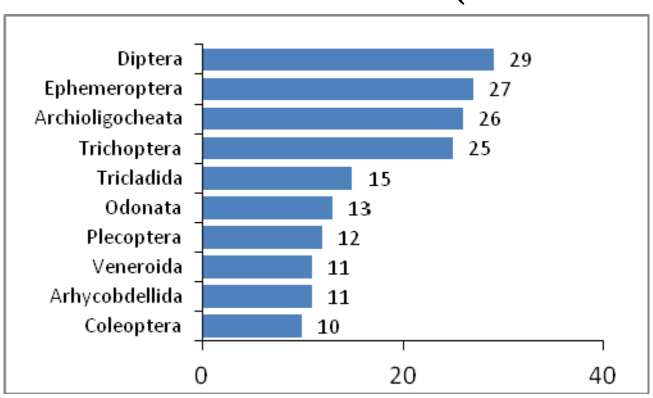


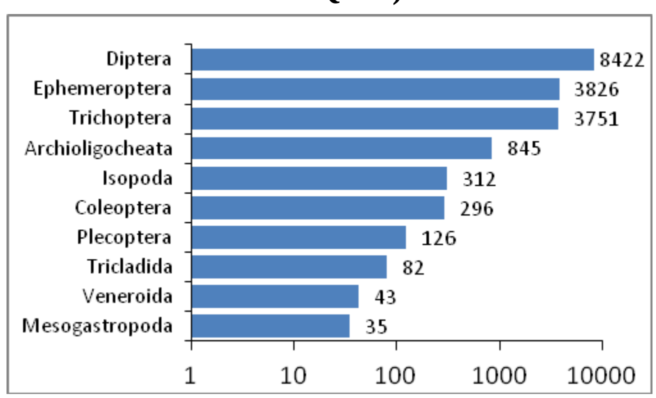


Site occurrence frequency

Relative abundance

COMB (combined data of KNAEMP and MDEQP51)





Site occurrence frequency

Relative abundance

Figure 5.2. Site occurrence frequencies and relative abundances of the top ten macroinvertebrate taxa by KNAEMP, MDEQP51, and COMB (combined data). Left column has site occurrence frequencies of each taxa and right column has relative abundances of each taxa. Order taxa were used for class Insecta and class taxa were used for others.

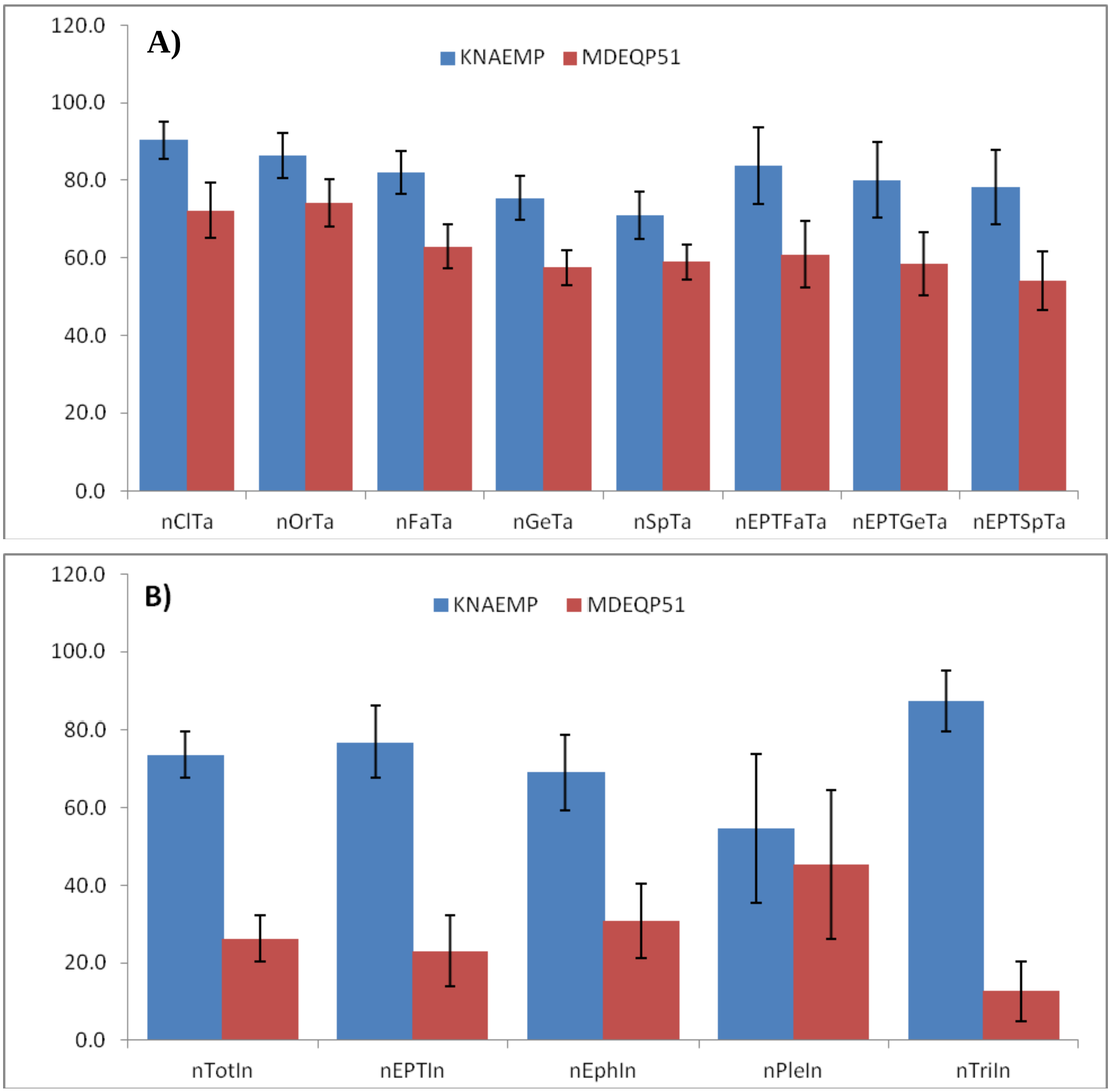


Figure 5.3. Mean macroinvertebrate sampling methodology efficiency (MSME) with 95 percentile ranges for KNAEMP and MDEQP51.

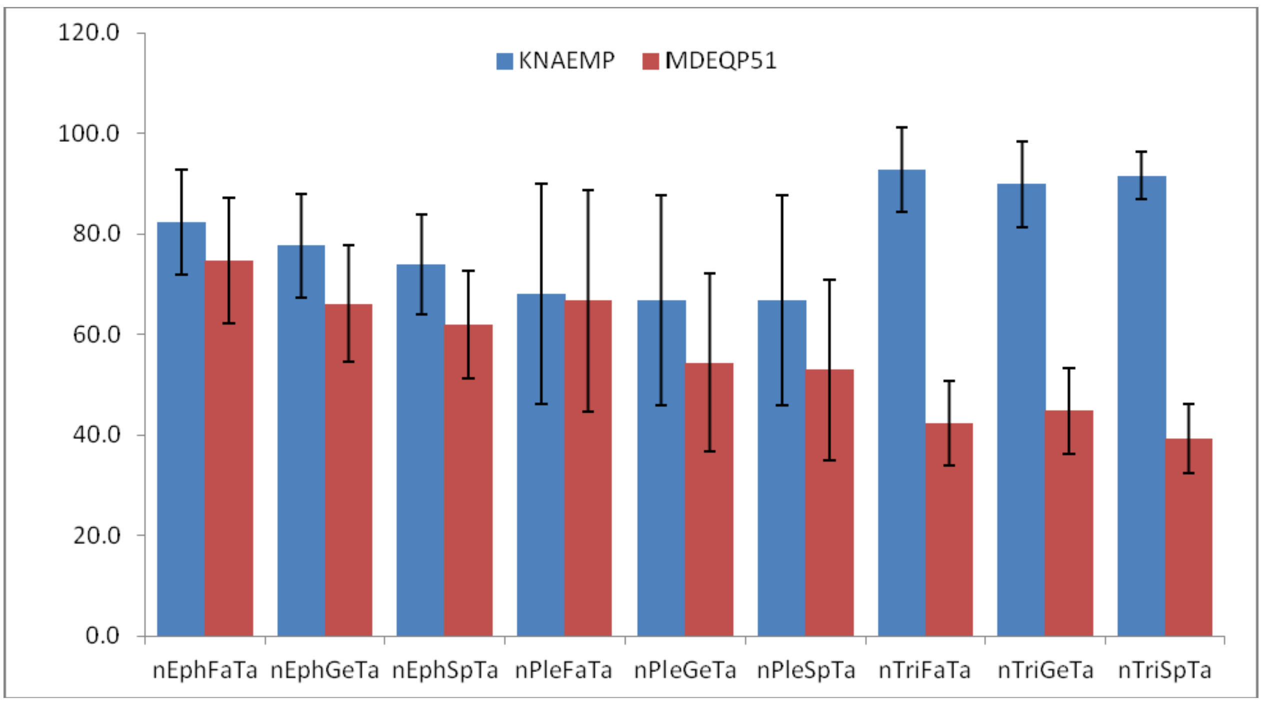


Figure 5.4. Mean macroinvertebrate sampling methodology efficiency (MSME) with 95 percentile ranges for KNAEMP and MDEQP51.

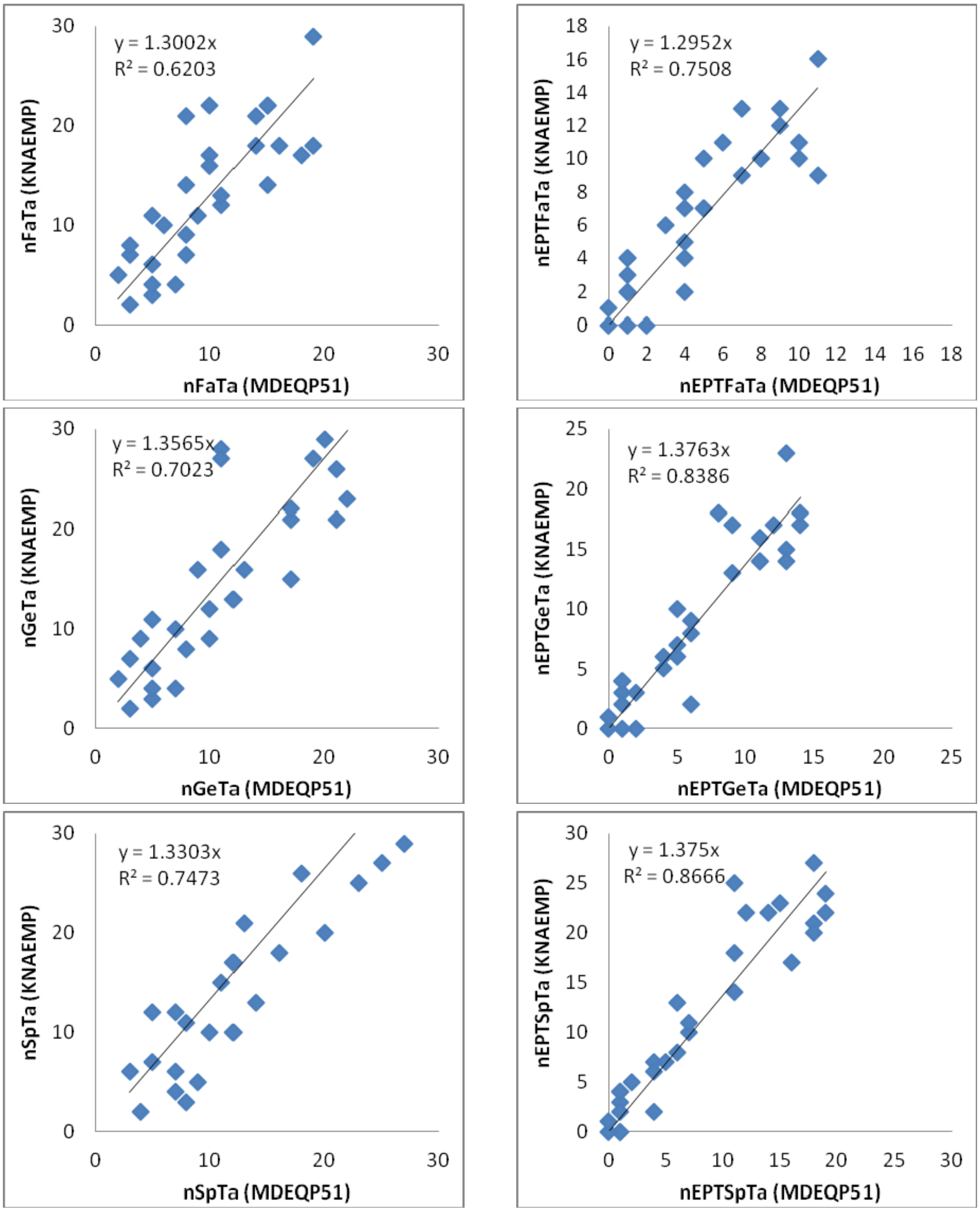


Figure 5.5. Comparison of taxa richness of each metric between KNAEMP and MDEQP51 and correction equations for sampling method bias.

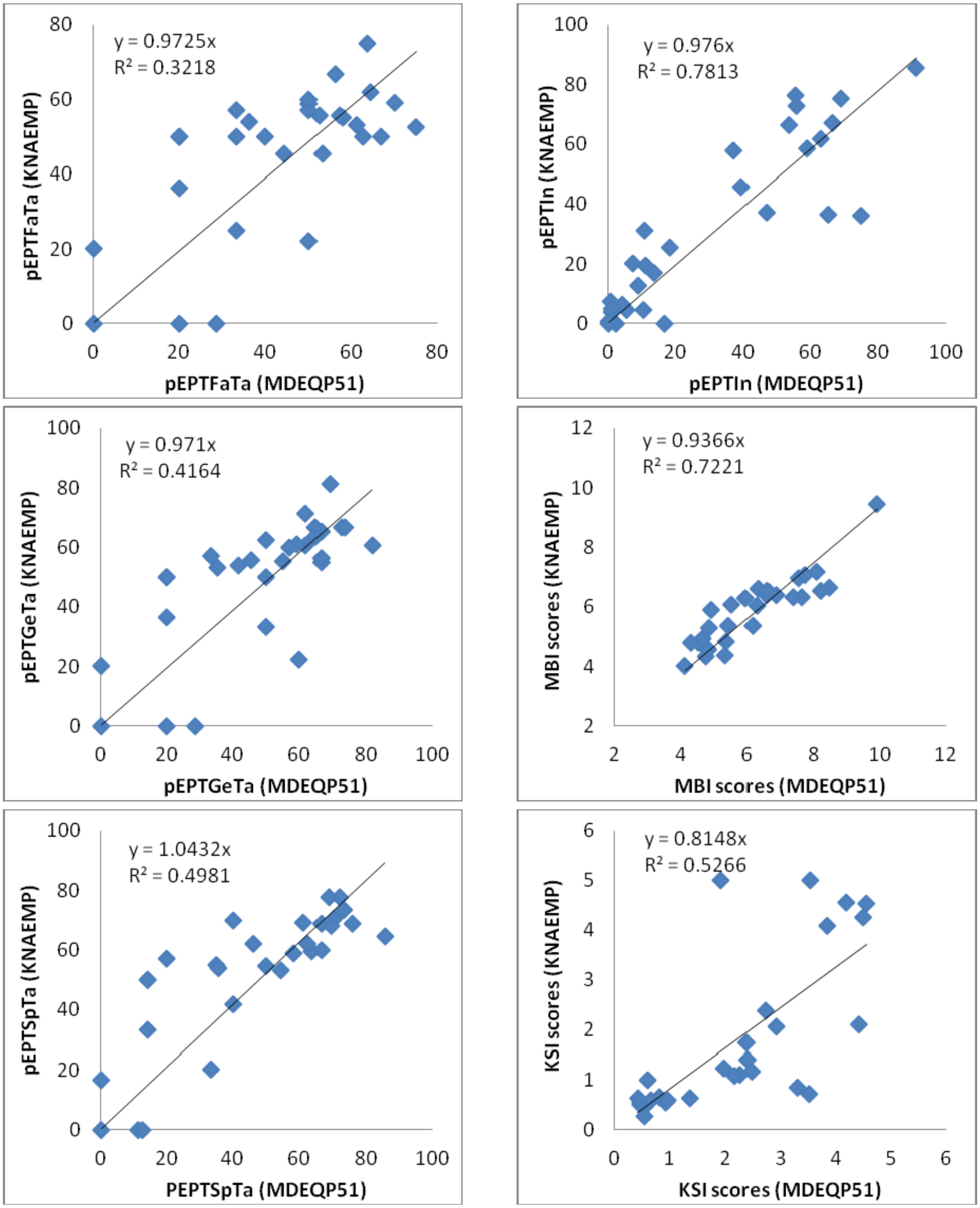


Figure 5.6. Comparison of benthic macroinvertebrate metric values between KNAEMP and MDEQP51 and correction equations for sampling method bias.

Before standardization **After standardization**

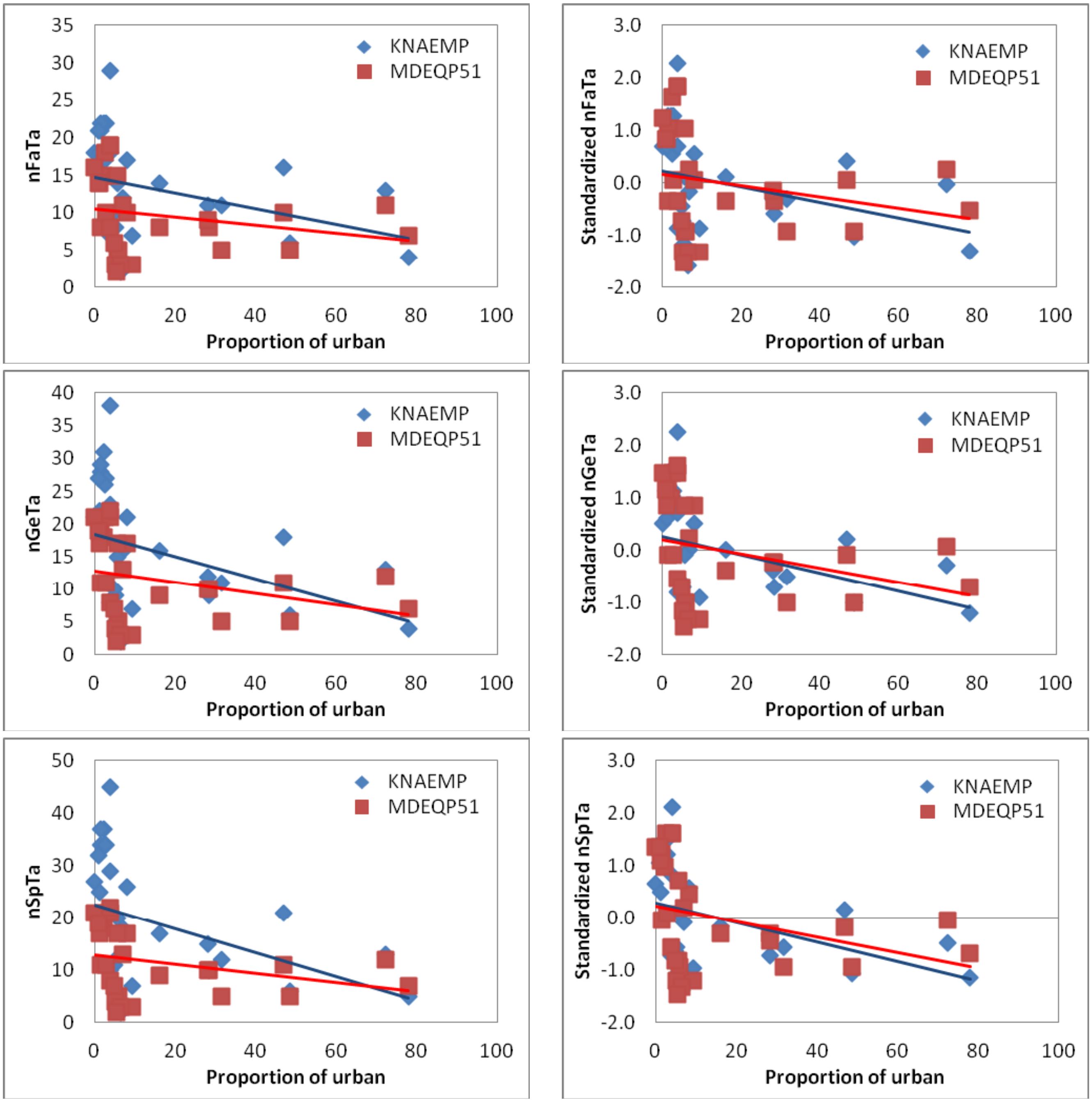


Figure 5.7. Comparison of taxa richness of two biological monitoring programs against proportion of urban landuse for each taxonomic resolution. Graphs on the left columns show data set before the standardization of taxa richness and graphs on the right columns show data set after the standardization of taxa richness.

Before standardization **After standardization**

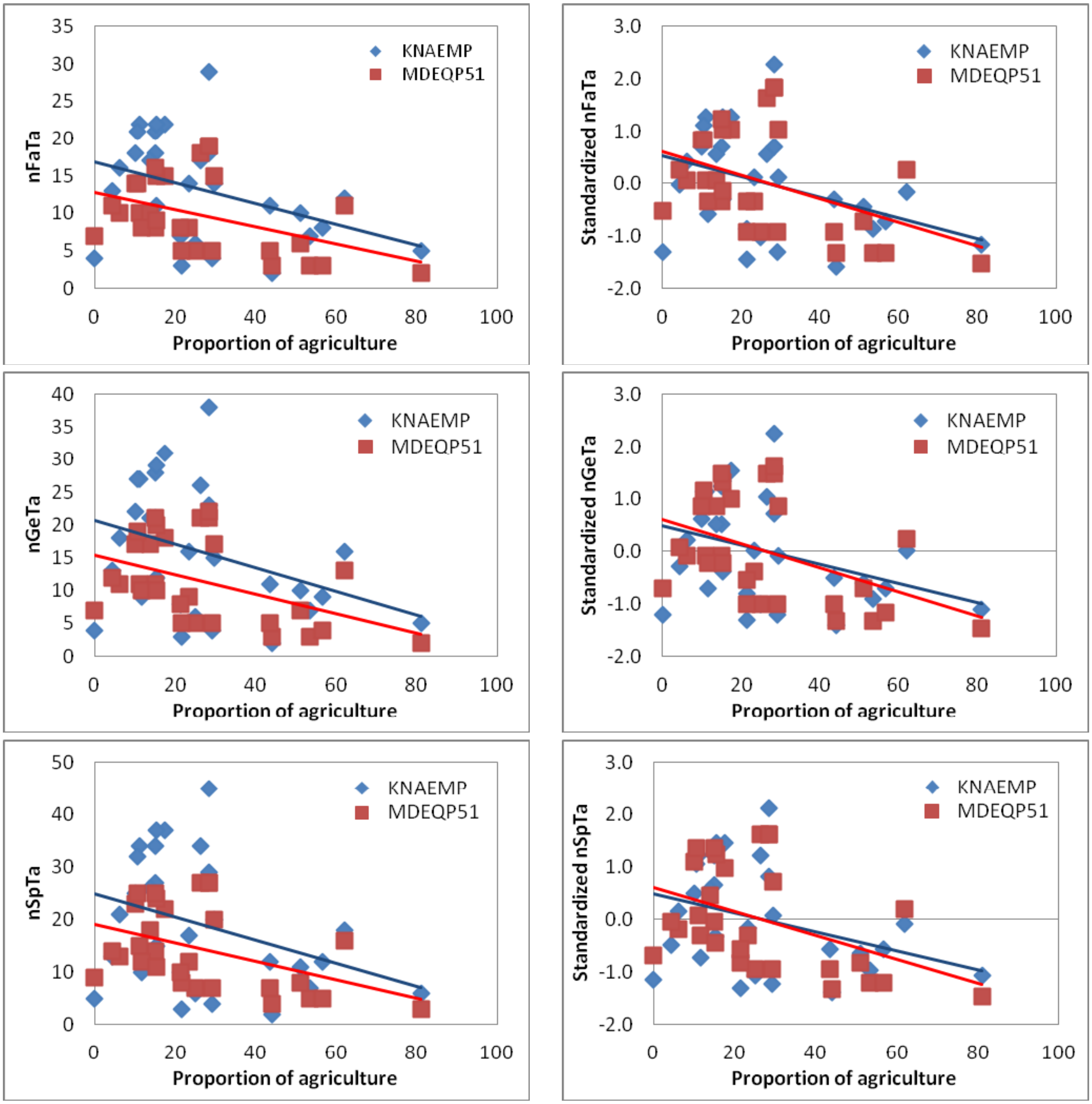


Figure 5.8. Comparison of taxa richness of two biological monitoring programs against proportion of agricultural landuse for each taxonomic resolution. Graphs on the left columns show data set before the standardization of taxa richness and graphs on the right columns show data set after the standardization of taxa richness.

Literature cited

Allan, J. D. and M. M. Castillo. 2007. Stream ecology: structure and function of running waters. Second edition. Springer. Dordrecht, The Netherlands.

Allan, J. D., P. B. McIntyre, S. D. P. Smith, B. S. Halpern, G. L. Boyer, A. Buchsbaum, G. A. Burton Jr., L. M. Campbell, W. L. Chadderton, J. J. H. Ciborowski, P. J. Doran, T. Eder, D. M. Infante, L. B. Johnson, C. A. Joseph, A. L. Marino, A. Prusevich, J. Read, J. B. Rose, E. S. Rutherford, S. P. Sowa, A. D. Steinman. 2013. Joint analysis of stressors and ecosystem services to enhance restoration effectiveness. Proceedings of the National Academy of Sciences. 110 (1): 372-377.

Baker, E. A., K. E. Wehrly, P. W. Seelbach, L. Wang, M. J. Wiley, and T. Simon. 2005. A multimetric assessment of stream condition in the Northern Lakes and Forests ecoregion using spatially explicit statistical modeling and regional normalization. Transactions of the American Fisheries Society. 134 (3): 697-710.

Barbour, M. T. and J. Gerritsen. 1996. Subsampling of benthic samples: a defense of the fixed-count method. Journal of the North American Benthological Society. 15 (3): 386-391.

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

Blalock, H. M. 1970. Estimating measurement error using multiple indicators and several points in time. American Sociological Review. 35 (1): 101-111.

Blocksom, K. A., J. P. Kurtenbach, D. J. Klemm, F. A. Fulk, and S. M. Cormier. 2002. Development and evaluation of the lake macroinvertebrate integrity index (LMII) for New Jersey lakes and reservoirs. Environmental Monitoring and Assessment. 77: 311-333.

Bonar, S. A. and W. A. Hubert. 2002. Standard sampling of inland fish: benefits, challenges, and a call for action. Fisheries. 27: 10-16.

Cao, Y., C. P. Hawkins, and A. W. Storey. 2005. A method for measuring the comparability of different sampling methods used in biological surveys: implications for data integration and synthesis. Freshwater Biology. 50: 1105-1115.

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

Clarke, R. T., J. Davy-Bowker, L. Sandin, N. Friberg, R. K. Johnson, and B. Bis. 2006. Estimates and comparisons of the effects of sampling variation using ‘national’ macroinvertebrate sampling protocols on the precision of metrics used to assess ecological status. Hydrobiologia. 566: 477-503.

Clarke, R. T. and D. Hering. 2006. Errors and uncertainty in bioassessment methods – major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia. 566: 433-439.

Cummins, K. W. 1962. An evaluation of some techniques for the collection and analysis of benthic samples with special emphasis on lotic waters. American Midland Naturalist. 67: 477-504.

Dahl, Joakim, and R. K. Johnson. 2004. A multimetric macroinvertebrate index for detecting organic pollution of streams in southern Sweden. Archiv Fur Hydrobiologie. 160 (4): 487-513.

Diamond, J. M., M. T. Barbour, and J. B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. Journal of North American Benthological Society. 15 (4): 713-727.

Diana, J. S. 2004. Biology and ecology of fishes, Second edition. Cooper Publishing Group. Traverse City, MI.

Doberstein, C. P., J. R. Karr, and L. L. Conquest. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. Freshwater Biology. 44: 355-371.

Dodds, W. K. 2006. Eutrophication and trophic state in rivers and streams. Limnology and Oceanography. 51: 671–680.

ESRI. 2009. ArcGIS. Version 9.1. (Build 722), Redlands, CA.

Freeman, B. J., H. S. Greening, and J. D. Oliver. 1984. Comparison of three methods for sampling fishes and macroinvertebrates in a vegetated wetland. Faculty Publications and Presentations. Paper 89.

Friberg, N., L. Sandin, M. T. Furse, S. E. Larsen, R. T. Clarke, and P. Haase. 2006. Comparison of macroinvertebrate sampling methods in Europe. Hydrobiologia. 566: 365-378.

Furse, M., D. Hering, O. Moog, P. Verdonschot, R. K. Johnson, K. Brabec, K. Gritzalis, A. Buffagni, P. Pinto, N. Friberg, J. Murray-Bligh, J. Kokes, R. Alber, P. Usseglio-Polatera, P. Haase, R. Sweeting, B. Bis, K. Szoszkiewicz, H. Soszka, G. Springe, F. Sporka, and I. Krno. 2006. The STAR project: context, objectives, and approaches. Hydrobiologia. 566: 3-29.

Gerth, W. J. and A. T. Herlihy. 2006. Effects of sampling different habitat types in regional macroinvertebrate bioassessment surveys. Journal of North American Benthological Society. 25 (2): 501-512.

Hellawell, J. M. 1986. Biological indicators of freshwater pollution and environmental management. Pollution Monitoring Series. Elsevier Applied Science. London, New York.

Hilsenhoff, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20: 31-39.

Houston, L., M. T., Barbour, D. Lenat, and D. Penrose. 2002. A multi-agency comparison of aquatic macroinvertebrate-based stream bioassessment methodologies. Ecological Indicators. 1: 279-292.

ITFM (Intergovernmental Task Force on Monitoring Water Quality). 1995. The strategy for improving water – quality monitoring in the United States. Final report of the Intergovernmental Task Force on Monitoring Water Quality. Office of Water Data Coordination, U. S. Geological Survey, Reston, Virginia, Open-File Report 95-742.

Jun, Y., N. Kim, S. Kwon, S. Han, I. Hwang, J. Park, D. Won, M. Byun, H. Kong, J. Lee, and S. Hwang. 2011. Effects of land use on benthic macroinvertebrate communities: Comparison of two mountain streams in Korea. Annales de Limnologie – International Journal of Limnology. 47: S35-S49.

Karr, J. R. 1981. Assessment of biotic integrity using fish communites. Fisheries. 66: 21-27.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: Amethod and its rationale. Special publication 5. Illinois Natural History Survey.

King, R. S. and C. J. Richardson. 2002. Evaluating subsampling approaches and macroinvertebrate taxonomic resolution for wetland bioassessment. Journal of the North American Benthological Society. 21 (1): 150-171.

Klemm, D. J., K. A. Blocksom, and F. A. Fulk. Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing Mid-Atlantic highlands streams. 2003. Environmental management. 31 (5): 656-669.

Kwon, O. K., D. K. Min, J. R. Lee, J. S. Lee, J. G. Je, and B. L. Choe. 2001. Korean mollusks with color illustration (In Korean). Hanguel Publishing Company, Seoul, Korea.

Larimore, R. W. 1961. Fish population and electrofishing success in a warm-water stream. Journal of Wildlife Management. 25: 1-12.

Launois, L., J. Veslot, P. Irz and C. Argillier. 2011. Development of a fish-based index (FBI) of biotic integrity for French lakes using the hind casting approach. Ecological Indicators. 11: 1572-1583.

Merritt, R. W., and K. W. Cummins. 1996. An introduction to the aquatic insects of North America. Kendall Hunt Publishing, Dubuque, Iowa.

MDEQ (Michigan Department of Environmental Quality). 1997. Qualtitative biological and habitat survey protocols for wadeable streams and rivers. Michigan Department Environmental Quality (MDEQ), Surface Water Quality Division, Great Lakes and Environmental Assessment Section, GLEAS Procedure 51 (Revised 2002). Lansing, Michigan.

Michigan Department of Natural Resources (MDNR). 2002. Huron River Plan (Livingston, Washtenaw, Oakland, Monroe and Wayne Counties). Michigan Department of Natural Resources, Fisheries Division. Lansing, Michigan.

MEA (Millennium Ecosystem Assessment). 2005. The millennium ecosystem assessment series. Ecosystems and human well-being: Multiscale assessment. Vol. 4. Island Press. NW, Washington DC.

Moulton, S. R., J. G. Kennen, R. M. Goldstein and J. A. Hambrook. 2002. Revised protocols for sampling algal, invertebrate and fish communities as part of the National Water-Quality Assessment Program. USGS Open-File Report 02-150. Reston, Virginia.

NIER (National Institute of Environmental Research). 2009. Investigation and evaluation of aquatic ecosystem of South Korean streams and rivers. Final Report. May 2009. Ministry of Environment. S. Korea.

Ohio EPA (Ohio Environmental Protection Agency. 1987. Biological criteria for the protection of aquatic life: volumes I-III. Ohio Environmental Protection Agency, Columbus, Ohio.

Park, K. S. 2007. Effects of different sampling effort and taxonomic resolution on assessment metrics. Masters thesis. University of Michigan. Ann Arbor, MI.

Parsons, M. and R. H. Norris. 1996. The effects of habitat-specific sampling on biological assessment of water quality using a predictive model. Freshwater Biology. 36: 419-434.

Rinella, D. J. and J. W. Feminella. 2005. Comparison of benthic macroinvertebrates colonizing sand, wood, and artificial substrates in a low-gradient stream. Journal of Freshwater Ecology. 20 (2): 209-220.

Riseng, C. M., M. J. Wiley, R. J. Stevenson, and P. J. Seelbach. 2006. Comparison of coarse versus fine scale sampling on statistical modeling of landscape effects on fish communities of the Muskegon River. Pages 555-576 in R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Landscape influences on stream habitats and biological assemblages. American Fisheries Society, Symposium 48, Bethesda, Maryland.

Riseng, C. M., M. J. Wiley, P. W. Seelbach, and R. J. Stevensen. 2010. An ecological assessment of Great Lakes tributaries in the Michigan Peninsulas. Journal of Great Lakes Research. 36: 505-519.

Riseng, C. M., M. J. Wiley, R. W. Black, and M. D. Munn. 2011. Impacts of agricultural land use on biological integrity: a causal analysis. Ecological Applications. 21 (8): 3128-3146.

Sladecek, V. 1973. System of water quality from the biological point of view. Archiv für Hydrobiologia. 566: 543-555.

Smith, D. G. 2001. Pennak's freshwater invertebrates of the United States – Porifera to Crustacea. John Wiley and Sons, New York.

Šporka, F. H. E. Vlek, E. Bulánková, and I. Krno. 2006. Influence of seasonal variation on bioassessment of streams using macroinvertebrates. Hydrobiologia. 566: 543-555.

SPSS, Inc. 2003. SPSS for Windows Version 12.0. SPSS, Inc., Chicago, Illinois.

Stark. J. D. 1993. Performance of the macroinvertebrate community index: effects of smapling method, sample replication, water depth, current velocity, and substratum on index values. New Zealand Journal of Marine and Freshwater Research. 27 (4): 463-478.

Strahler, A. N. 1957. Quantitative analysis of watershed geomorphology. Transactions of the American Geophysical Union. 38 (6): 913-920.

U. S. Environmental Protection Agency (EPA). 2006. The evaluation of methods for creating defensible, repeatable, objective and accurate tolerance values for aquatic taxa. EPA/600/R-06/045. Office of Research and Development, National Exposure Research laboratory, Cincinnati, OH.

US EPA (U. S. Environmental Protection Agency). 2007. National water quality inventory: report to congress. EPA/841/R-07/001. Office of Water, Washington, DC.

Velleman, P. F. 1997. DataDesk Version 6.0: Statistics Guide. Data Description. Ithaka, NY.

Walsh, C. J. 1997. A multivariate method for determining optimal subsample size in the analysis of macroinvertebrate samples. Marine and Freshwater Research. 48: 241-248.

WAMIS (Water Management Information System). 2013. The Ministry of Land, Infrastructure, and Transport. Stream flow gauges and flow data. Republic of Korea. Available: Http://wamis.go.kr.

Wang , L., J. Lyons, P. Kanehl, and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28: 255-266.

Wang, L., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela, and P. Stewart. 2003. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest ecoregion, USA. Canadian Journal of Fisheries and Aquatic Sciences 60: 491-505.

Wiley, M. J., P. W. Seelbach, K. Wehrly, and J. Martin. 2003. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. Chapter 12 in: T. P. Simon, Editor. Biological response signatures: indicator patterns using aquatic communities. CRC Press, Boca Raton, Florida.

Won, D. H., S. J. Kwon, and Y. C. Jun. 2005. Aquatic insects of Korea (In Korean). Korea Ecosystem Service Publishing Company. Seoul, Korea.

Won, D. H., Y. C. Jun, S. J. Kwon, S. J. Hwang, K. G. Ahn, and J. K. Lee. 2006. Development of Korean Saprobic Index using benthic macroinvertebrates and its application to biological stream environment assessment. Journal of Korean Society on Water Quality. 22: 768-783.

Yoon, I. B. 1995. An illustration of aquatic insects in Korea (In Korean), Jeonghaengsa Publishing Company. Seoul, Korea.

Zelinka, M. and P. Marvan. 1961. Zur präzisierung der biologische klassifikation der reinheit fliessender gewässer. Archiv fur Hydrobiologie. 57: 389-407.

Chapter 6 : Meta-analysis with ecological assessment data: a case study of Michigan and South Korean streams

Abstract

In this study I examine issues of comparability in ecological assessment using a case study of stream assessment data from Michigan and South Korea. Initial comparisons of biological and landscape data indicated that direct comparisons of rapid bioassessment survey results would be difficult due to differences in sampling methods, differing sets of ecological covariates, and suspected differences in the intensity of anthropogenic stresses. Methodological biases in the data were identified and corrected using gear calibration functions (chapters 4 and 5). Regional normalization (residualization) corrected statistically significant biases in observed land use stressor-response relationships from both regions. Normalized multimetrics indicated that in both regions, fish and invertebrate communities were more sensitive to urban land use than to agricultural land use; and that S. Korean streams were more seriously degraded than Michigan streams. LU stressor-response relationships for fish varied significantly between regions but not for benthic macroinvertebrates. This difference in response may reflect distinct zoogeographic histories of the two regions since taxonomic similarity is high for the aquatic insect fauna but relatively low for the fish fauna. A deeper understanding of regional biases in assessment data sets and methodologies is essential to inter-regional and global evaluations of anthropogenic impact and to the successful transfer of assessment tools and technologies to the developing world.

Introduction

Over the past four decades anthropogenic alterations of natural communities and ecosystems have been increasing concern for governments everywhere. As a result biological assessment data (community composition data used to infer ecological condition) has become increasingly available in many parts of the world. At the same time awareness of the global nature of ecological change drives growing interest in larger-scale regional and even global assessments of ecological condition, and inter-regional transfers of environmental assessment, planning, and control tools and technologies. For example, the Millennium Ecosystem Assessment (MEA 2005) by the United Nations Environment Programme (UNEP) has recently evaluated world biodiversity using a biological indicator (total taxa richness) across 33 global sub-regions. The European Union Water Framework Directive (EU-WFD) assessed European streams and rivers using a newly developed standard protocol; and compared its accuracy to the (eleven) existing national rapid bioassessment programs (RBPs) (see Table 6.1 for definitions of acronyms frequently used in this chapter) currently used by EU member nations (Furse et al. 2006). Similar but separate transnational studies in Europe have also been carried out for lakes (Lanois et al. 2011, Argilliar et al. 2013). The US Environmental Protection Agency (US EPA) both through state-developed annual assessment reporting and its national surveys Environmental Monitoring and Assessment Program (EMAP) is evaluating the status and trends of national ecological resources in 3 major and 9 ecological regions for the United States (US EPA 2007).

However, each of these efforts experienced substantive difficulty in integrating existing assessment datasets from across their focal regions, and all ended up requiring new (and redundant from a public policy perspective) data collections using new standardized methodologies, or developing new standardized indicator metrics to be applied in the larger scale analysis. Difficulties comparing interregional and international assessment datasets arise for a variety of reasons. Most assessment methodologies and methods have been developed for limited geographic areas (typically civil units) to maximize both information gleaned and spatial coverage under substantial budget constraints. Choices of collection gear employed, types of organisms collected, taxonomic resolution, size of area sampled, etc. often vary widely between even adjacent states or countries (Bryce and Clarke 1996, US EPA 2007). But even when protocols and effort allocations are identical, natural geographic/spatial variation in ecological communities and processes can lead to measurement biases and inconsistencies; especially between biogeographically different regions (Diamond et al. 1996, Riseng et al. 2011).

Questions of comparability, however, are not restricted to larger-scale assessments. Many regional bioassessment programs use multiple indicator variables to provide some control over individual metric biases; this approach is explicitly incorporated into common assessment multi-metric indicators (MMI) like the Index of Biotic Integrity (Karr and Chu 2000). But differences in indicator responses to the same stressor gradients can also reflect problematic scaling or methodological biases that obscure assessment results. For example South Korea's Nationwide Aquatic Ecological Monitoring Program (KNAEMP, NIER 2009) uses two MMI indicators, one based on fish community sampling (the Korean Index of Biotic Integrity; KIBI) modeled on Karr's IBI (Karr 1981), the other based on macroinverebrate sampling (Korean Saprobic Index; KSI) based on the saprobic valency concept (Zelinka and Marvan 1961). Results of the national assessments varied widely between these MMIs (Figure 6.1), resulting in considerable controversy both within and outside the government (NIER 2009).

Problems related to comparability of assessment datasets and the metrics derived from them commonly involve either issues of differing scaling or measurement bias. Since assessment metrics are usually interpreted in context of deviations from a specified criterion, bias can arise in measurement of the metric value itself, or in specification of the reference condition for the metric by which it is interpreted, or both (Wiley et al. 2003). Biases can be methodological (see chapters 4 and 5, this dissertation); but they can also be statistical if they arise from covariances among natural driving variables (e.g. hydrogeology, river network structure, catchment size, and land use patterns) which in turn influence the spatial patterning of biological communities (Wiley et al. 2003, Baker et al. 2005, Schoolmaster et al. 2013). Covariance relationships result in complex joint probability distributions which require experimental or statistical control to yield unbiased estimates of response in dependent variables (Pearl 2009, Schoolmaster et al. 2013). Both methodological and statistical biases can lead to a failure to correctly diagnose ecological status and interpret empirical stressor-response relationships; potentially resulting in inappropriate regulatory policies and management actions.

This dissertation focused on issues of data comparability and integrability in the context of RBPs. In this last chapter I compared ecological assessment data from S. Korea and Michigan, two geographically disparate regions, in a case study format. Ecologically, both regions lie in the same ‘temperate seasonal forest’ biome (Ricklefs 2008); however they have clear differences in patterns of seasonal air temperature and precipitation, land use (LU), geology, and hydrology (see Chapters 1, 2, and 3). Differences in historical biogeography that have produced different taxa composition in the two regions also complicate data comparison and interpretation. Furthermore the two regions are characterized by very different human population densities (485.6 people /km^2 for S. Korea (2010 summary; KOSIS 2015) and 67.5 people /km^2 for

Michigan (2010 summary; US Census Bureau 2015)) and cultural practices with respect to LU management. These differences may be sufficient to produce fundamental differences between Michigan and S. Korean LU stressor-response relationships, which would further complicate comparisons of ecological status.

My specific objectives are to 1) compare Korean and Michigan ecological datasets, 2) explore the impacts of known sampling biases (Chapters 4 and 5) and regionally covarying landscape properties (Chapters 2 and 3) on their respective LU stressor-response relationships, and 3) determine the extent to which explicit corrections for methodological and statistical biases lead to altered interpretations of assessment results. Finally, this case study will 4) address whether underlying LU stressor-response relationships and rates of impairment vary between the two regions.

Materials and methods

Case study approach

Fish and benthic macroinvertebrate sample data used in this study were obtained from government sponsored rapid bioassessment programs (RBPs) in Michigan and South Korea. I focus on these regions to explore issues of comparability of data in meta-analyses aimed at integrated and/or comparative assessment. Together they provide a useful platform from which to develop a case study highlighting typical problems (and potential solutions) that arise in comparing bioassessment data from differing geographic contexts.

Biological data from Michigan and South Korean regions were collected by government-supported crews using their own state or national standard RBPs. Biological data from Michigan (n = 803) were collected using methods specified as RAP Procedure 51 (MDEQ 1997, Park 2007, also see Riseng et al. 2008). Biological data from South Korea (n = 684) were sampled as a part of the KNAEMP of the Ministry of Environment, South Korea and followed their own and quite different RAP guidelines (NIER 2009, Hwang et al. 2011, Jun et al. 2011, Lee et al. 2011, Yoon et al. 2011). The goal of both programs is the assessment of ecological condition in their respective regions. In this case study I use these datasets to address two larger-scale questions requiring comparisons of assessment data from both sources:

1. Do underlying LU stressor-response relationships vary between the two regions?
2. How similar is the extent of ecological impairment in S. Korea and Michigan?

The first question is implicit in the second, and both require a careful analysis of differences in methodology, ecological context, metric selection and interpretation. Building on work described in previous chapters, I will approach the questions with three different levels of data set preparation. "*Raw*" data and metrics will refer throughout to the original regional datasets as collected and processed using the regional protocols described below. *"Method corrected"* data will be used to refer to transformed data sets in which statistically derived calibration functions (developed in chapters 4 and 5) are used to adjust sample counts to reflect sampling biases inherent in the methodology of the two regions. "*Normalized*" data and metrics will refer to metrics developed using a process of regional normalization described below which both re-centers and re-scales the original assessment datasets.

Fish sampling methods

Fish assemblage data from the two regions were collected using RBPs which employed different sampling methodologies and indicator metrics. However, the main goal of both protocols is to provide a representative species list and reasonable estimates of relative abundance (MDEQ 1997, NIER 2009). Protocols are summarized below, but more detailed descriptions are available in Chapter 4 (Calibration research for fish sampling data between Michigan and S. Korean regions) and elsewhere (MDEQ 1997, Wiley et al. 2003, Riseng et al. 2010 for Michigan protocol and An et al. 2005, NIER 2009, Hwang et al. 2011, Lee et al. 2011 for Korean protocol).

Sampling dates for Michigan fish data (n = 746) ranged from 1989 to 2004. Fish communities in wadeable streams were sampled using primarily single-pass DC electrofishing from downstream to upstream with either back-pack or tow-barge electrofisher with no block nets (MDEQ 1997). A back-pack electrofisher was mostly used for smaller streams. Sampling reach distance ranged from 30 to 90 meters over 30 minutes with a minimum goal of 100 fish per site.

Korean fish data (n = 684) were all collected in 2009 because it was the only data set available after the South Korean government initiated a standard nation-wide assessment program (NIER 2009). Fish sampling for S. Korean region was conducted by cast net (mesh size: 5 mm; net diameter: 6.5 m) with the combination of small hand seine (mesh size: 4 mm and net size: 1×1 m). The cast netting team consisted of two persons over approximately one hour; one with a cast net working through the whole reach, and the other person alternately using both a cast net and a hand seine (NIER 2009).

Benthic macroinvertebrate sampling methods

As with fish, benthic macroinvertebrate data from the two regions were collected using RBPs with different sampling methodologies and indicator metrics. Michigan (MDEQP51) samples characterize the structure of invertebrate communities in terms of relative abundances of taxa rather than absolute density (Moulton et al. 2002), whereas KNAEMP samples estimate the absolute density of invertebrate taxa (NIER 2009, Jun et al. 2011). Sampling protocols are summarized below, but more detailed descriptions are available in Chapter 5 (Calibration research for invertebrate sampling data between Michigan and Korean regions) and elsewhere (MDEQ 1997, Park 2007, Riseng et al. 2010 for Michigan protocol and NIER 2009, Jun et al. 2011 for Korean protocol).

Sampling dates for the Michigan dataset (n = 774) ranged from 1989 to 2004. Sampling of invertebrate assemblages was performed using D-frame dip nets (250 μm mesh size) for 30 minutes at each site by one person or by several persons (shares minutes by number of people in this case). Samples from all habitats were combined in a bucket and then 100 organisms were randomly selected from the composite sample for further analysis (Merritt and Cummins 1996, MDEQ 1997, Riseng et al. 2006). The 100 selected organisms were preserved in 70 % ethanol and returned to the laboratory for identification and enumeration (Merritt and Cummins 1996).

Korean invertebrate data (n = 684) collected in 2009 were used to match the analysis of Korean fish data. KNAEMP samples were quantitatively collected at riffle habitats using a surber sampler (30 cm × 30 cm, 1 mm mesh size). Three samples at each site were taken from randomly selected riffles in a designated stream reach and placed into a 500 ml plastic bottle after removing large substrate and debris. Then 70 % ethanol was added to preserve samples for further identification and enumeration of all animals collected.

Indicator metrics

In total thirteen indicator metrics were selected for this analysis based on assemblage composition, analytical assessment methodologies, and preliminary tests of indicator metrics in Chapters 4 and 5 (Table 6.2). Fish metrics (a total of seven) included: number of total fish species (nFiSp), number of intolerant fish species (nFiInt), percentage of tolerant fish individuals (pFiTol), percentage of omnivorous fish individuals (pFiOmn), percentage of insectivorous fish individuals (pFiIns), Korean Index of Biological Integrity (KIBI) score, and MDEQP51 (P51Fi) score for fish. These metrics were chosen to represent fish assemblage measures of species richness, tolerance guild, feeding guild, and biological integrity (Barbour et al. 1999, Wiley et al. 2003, Riseng et al. 2010). P51Fi and KIBI scores are ecological bioassessment multimetric-indices (MMIs) for Michigan (MDEQ 2002) and S. Korea (NIER 2009), respectively, and both aim to reflect the ecological health of a sampling site. These regional MMIs could only be calculated for their respective regions because they use different kinds of base metrics and scoring schemes to produce an assessment. P51Fi (MDEQ 2002) includes ten metrics and each metric score has three classes (+1 for excellent, 0 for acceptable, and -1 for poor condition). A fish score for a site is calculated based on the sum of each of ten metrics and ranges from +10 (Excellent condition) to -10 (poor condition). In contrast KIBI (NIER 2009, based on the IBI concept of Karr 1981) has eight multi-metrics and each metric score also has three classes (+5 for excellent, +3 for acceptable, and +1 for poor condition). A fish score for a site is calculated based on the sum of each of eight metrics and ranges from +40 (Excellent condition) to +8 (poor condition).

Benthic macroinvertebrate indicator metrics included: number of total invertebrate families (nFaInv), number of Ephemeroptera-Plecoptera-Trichoptera (EPT) families (nEPTFa), percentage of total individuals that were EPT (pEPTIn), Macroinvertebrate Biotic Index (MBI) score, Korean Saprobic Index (KSI) score, and MDEQP51 (P51Inv) score for benthic macroinvertebrates. These metrics are representative invertebrate assemblage measures of taxa richness, tolerance guild, and biological integrity (Barbour et al. 1999, Wiley et al. 2003, Riseng et al. 2010). P51Inv and KSI are multi-metric rapid bioassessment indices used in Michigan (MDEQ 2002) and S. Korea (NIER 2009), respectively. Each MMI produces an assessment score for their region only because of the different scoring schemes and taxonomic criteria to produce an assessment. P51Inv (MDEQ 2002) includes nine metrics and each metric score has three classes (+1 for excellent, 0 for acceptable, and -1 for poor condition). An invertebrate P51Inv score for a site is calculated based on the sum of each of ten metrics and ranges from +9 (excellent condition) to -9 (poor condition).

The invertebrate scores of KSI and MBI have a negative relationship with ecological integrity, which means higher values indicate poor ecological status and degradation in water quality. KSI is a modified invertebrate index of biological integrity from the saprobic valency concept (Zelinka and Marvan 1961) and is currently used for the ecological assessment of

Korean streams and rivers (NIER 2009). Saprobic values and weighting factors have been summarized for 100 major Korean invertebrate taxonomic groups. KSI scores range from 0 (excellent condition) to 5 (poor condition); the KSI score of each site was calculated by averaging sum of saprobic value and weighting factor of each taxon collected (Won et al. 2006 and NIER 2009). MBI is taken from Hilsenhoff or EPA established biotic index values (Hilsenhoff 1987 and US EPA 2007). A tolerance value for each taxon ranged from 0 to 10 and the average MBI score of each site was calculated by averaging sum of a published tolerance value for each taxon collected (Hilsenhoff 1987, Riseng et al. 2006).

Nine of the thirteen metrics used here were also calibrated for sampling biases related to sampling gear (fish) or methods (benthic macroinvertebrates). Statistical equations for sampling bias correction are summarized in Table 6.2 and more details are described in Chapters 4 (for fish metrics) and 5 (for invertebrate metrics). However, both raw and method-corrected data are used to examine the effects of sampling methodologies on stream assessment and in regional normalization models.

Environmental data

For each fish or invertebrate sampling site, landscape-scale variables were summarized (Table 6.3, Appendix 6.1) by site or by catchment as appropriate using ArcGIS 9.1 (Brenden et al. 2006, ESRI 2009). Landscape-scale variables expected to influence natural stream biological assemblages included drainage area, water temperature, stream flow yields, and site slope. Landscape-scale stressors expected to influence biological assemblages included percent of urban and agriculture LUs in catchment and number of dams (Wiley et al. 2003, Riseng et al. 2006, Riseng et al. 2010).

The Michigan environmental data were summarized previously (Riseng et al. 2006); catchment boundaries of each site were delineated by the Michigan Department of Natural Resources (MDNR) from United States Geological Service (USGS) 1:100,000 scale topographic maps. Major LU categories (urban, agriculture, forest, forested wetland, nonforested wetland, and water) were summarized by catchment using 1998 IFMAP (raster-based) LU coverage (Brenden et al. 2006). Site slope was measured with the digital elevation map and number of dams with storage was summarized for each site's catchment based on the MDNR dam database. Field measurements of channel morphology (average wetted width and average wetted depth) and water temperature were included in the MDEQ data set. Stream flow yields (high (Q10Y), median (Q50Y), and low (Q90Y)) were collected by the USGS or modeled from landscape data (Seelbach et al. 2002).

For the S. Korean environmental data, catchment boundaries of each site were delineated using a watershed boundary map from the WAter Management Information System (WAMIS 2011), S. Korea. Because drainage areas of 51 Korean sites were much larger than the maximum site drainage area of Michigan region (Figure 6.2, Table 6.3, Appendix 6.1), I used two sets of Korean sites in many analyses. These were smaller but more directly comparable set (n = 633 sites) with larger sites (larger than 3,500 km^2) removed (Table 6.3); and the full set of Korean sites (n = 684 sites) which included the 51 largest river sites (Appendix 6.1). Major LU categories (urban, agriculture, forest, forested wetland, nonforested wetland, and water) were summarized by catchment using 2000 LU cover mapping from satellite image data (WAMIS 2011). I determined site slope from digital elevation maps, and the number of dams with storage was summarized for each site's catchment based on the WAMIS dam database. Field measurements of channel morphology (average wetted width and average wetted depth) and water temperature were included in the NIER data set (NIER 2009). Stream flow yields (high (Q10Y), median (Q50Y), and low (Q90Y)) were estimated by regression modeling as described in Chapter 2 (Estimation and classification of Korean stream flows) using datasets from the WAMIS (WAMIS 2011) and Korea Meteorological Administration (KMA 2011).

Region-specific ecological normalization of assessment metrics

Regional ecological normalization (Wiley et al. 2003, Baker et al. 2005, *aka* hindcasting in Kilgour and Stansfield 2006 and Argillier et al. 2013, dirty model in Hawkins et al 2010, whole set residualization (WSR) in Schoolmaster et al. 2013) was employed to compare LU stressor-response relationships of two regions and overall impairment of streams and rivers. Using this approach "normalized" MMIs were 1) re-centered on modeled site-specific reference conditions to correct for unintended but statistically detectable biases related to sampling or causal covariates (Seelbach et al. 2002, Wiley et al. 2003, Hawkins et al. 2010, Schoolmaster et al. 2013), and 2) re-scaled to reflect estimated regional variability (Kilgour and Stansfield 2006). To predict site-specific current condition (least disturbed condition; Davis and Henderson 1978, Zonneveld 1994, Wiley et al. 2003), multiple linear regression (MLR) models of metric scores were developed for both regions (Appendices 6.2-6.5). In MLR model construction I used a systematic manual stepwise progression to enter independent variables. Each MLR model included both natural variables associated with changes in community composition (drainage area, water temperature, site slope, and stream flow) and anthropogenic stressors related to human impacts (urban and agricultural land uses and number of dams). Variables were included in models if they were statistically significant ($\alpha = 0.05$) and improved the model fit (R^2 values and F statistics). Most variables were natural logarithm transformed [$\ln(x + c)$; c= 0, 0.001, or 1] to improve normality and linear relationships. However, MMIs (KIBI, KSI, P51Fi, and P51Inv) were not transformed since the raw data led to better model fits.

These MLR models were then used to estimate reference condition (i.e. hindcast reference, Kilgour and Stansfield 2006, Argillier et al. 2013; or reference condition model, Hawkins et al. 2010) at every site for each metric by setting anthropogenic stressor variables (urban LU, agricultural LU, and number of dams) to zero. Deviation values for each indicator metric were calculated by subtracting the predicted reference value from the appropriately transformed observed value. For pFiTol, MBI and KSI, the deviation values were calculated by subtracting the observed value from the expected value for each site since an increase in those metrics indicates a decline in ecological condition. Finally, the deviation values were scaled by dividing the computed deviation by the standard deviation of the modeled reference expectation to produce a normalized score scaled in standard deviation units. *Regionally normalized* in this context refers to the fact that the reference condition was predicted from a regional dataset and reflects a regional average expectation given the specific natural characteristics of the site (Riseng et al. 2010). The value of the normalized score typically ranges from -4 to +2 standard deviation units around the expected deviation value of zero. For example, a normalized score of zero implies that the metric value at a site is exactly as the regional reference model predicts or there is no evidence of adverse impact. A score of -2 would imply that the metric value is approximately 2 standard deviations below the reference condition for that region.

A normalized average MMI for fish and benthic macroinvertebrates was constructed by averaging a standard set of normalized indicator metrics in order to have a multimetric index score which could be computed for both Korean and Michigan sites (Wiley et al. 2003). My normalized composite fish MMI (CompFi) score included four indicator metrics: nFiSp, nFiInt, pFiTol, and pFiIns. My normalized composite invertebrate MMI (CompInv) score also included four indicator metrics: nFaInv, nEPTFa, pEPTIn, and MBI.

To summarize normalized composite MMIs I used a five level classification (five classes). Normalized scores above 0.5 were classed as “excellent,” scores between -0.5 and 0.5 as “good,” scores below -0.5 and above -1.0 as “fair,” scores below -1.0 and above -2.0 as “poor,” and scores below -2 as “very poor.” For the purposes of my analysis all sites classified as either “poor” or “ very poor” were considered to be biologically impaired; and sites classified as “excellent,” “good,” or “fair” were considered to be unimpaired.

Statistical analysis

Statistical summaries (mean, median, standard deviation, minimum, and maximum), Regression analyses, and scatter plots were conducted using Datadesk 6.0 (Velleman 1997). Independent samples t-test, GLM ANCOVA test, and Kolmogorov-Smirnov (K-S) test of data set were performed using SPSS 12.0 (SPSS, Inc. 2003).

Results

Differences in ecological context

The size of stream sites sampled (as indexed by drainage area) by the two programs was similar (Figure 6.2, Table 6.3, Appendix 6.1); reflecting the fact that both regions are peninsular and assessment sampling was restricted to wadeable sites. Frequency distributions differed somewhat between the two regions (K-S test, $p < 0.01$) due in part to the occurrence of a number of larger sites in South Korea which were sampled during low flow periods. The median drainage area for Michigan sites was 56 km^2 (Table 6.3), whereas the median for Korean sites was 141 km^2 (Appendix 6.1). Developed (urban and agricultural) LUs also differed between regions (Figures 6.3 and 6.4, Table 6.3, Appendix 6.1, K-S tests, $p < 0.01$). Although the frequency distribution of urban LU was largely similar, Korea had a number of sites with very high percentages of urban LU (as much as 100%), which did not occur in Michigan. The frequency distributions of agricultural LU were strikingly different. Many, if not most sites in Michigan had relatively high amounts of agricultural LU, whereas most Korean streams had lower amounts of agricultural LU in their catchment (Figures 6.3 and 6.4). Conversely Korean catchments were on average much more forested than Michigan catchments. ANCOVA tests showed that percentages of urban, agriculture, and forest differed significantly between regions ($p < 0.01$, Figure 6.4).

Hydrologically the rivers of the two regions differed in terms of flow and flow yields (Figure 6.5). At a similar size and exceedance frequency, Korean streams had on average higher flow rates and yields, reflecting both higher rainfall rates (chapter 2), higher catchment slopes

(Table 6.3, Appendix 6.1) and reduced permeability reflecting the mountainous terrain and shallow soils of the interior peninsula.

The biological communities of S. Korea and Michigan differed in important ways but also shared important similarities (see Table 1.4 in Chapter 1). Composition of the fish fauna had little overlap at the species level (percent similarity; 4.2% of Michigan species and 3.8% of Korean species), generic (similarity; 16.0% of Michigan genera and 14.0% of Korean genera), family or order levels (similarity; 42.9% and 50.0% of Michigan and 38.7% and 64.3% of Korean, respectively). In contrast the invertebrate faunas were more similar. Invertebrate data from MDEQP51 program were collected at the family level of identification, whereas Korean assessment program used species level identification; at the family level faunal similarity was 76.0% of Michigan families and 77.8% of Korean families; at the order level it was 80.0% of Michigan orders and 88.9% of Korean orders. Based on the assessment surveys, diversity is roughly comparable between the regions although methods of collecting and sampling, and species-area considerations complicate comparisons. Significant differences in taxa richness (Appendices 6.2-6.4) were observed for both fish and benthic macroinvertebrates for raw and method corrected data (Independent Samples t-test, $p< 0.05$). When taxa counts from raw data were compared, mean numbers of fish species and invertebrate families per sample (8.6 and 21.3, respectively) in Michigan were significantly higher than in S. Korean samples (7.5 and 10.7) (Appendix 6.2). The same pattern was seen when controlling for differences in stream size (Figure 6.2); richness in Michigan was significantly higher than in S. Korea, and this discrepancy increased with catchment area (Figure 6.6). However, when datasets were calibrated to compensate for the biases of sampling methods employed, mean differences in taxa richness for fish showed the reversed relationship from the raw data comparison (8.6 and 9.3 for Michigan and Korea, respectively; Appendix 6.3). Whereas the discrepancy between Michigan and Korea richness in invertebrate families was even more dramatic (27.7 versus 10.7; Michigan and Korea, respectively).

LU stressor-response relationships: individual indicator metrics

Based on the raw data calculations, individual fish and invertebrate indicator metrics were for the most part negatively and significantly correlated with urban and agricultural LUs in both the Korean and Michigan datasets (Table 6.4). Eight of nine indicators for each region were significantly correlated with agriculture; the exception being the pFiIns for Michigan and nFiSp for Korea (Table 6.4A). In Korea all indicators were also significantly correlated with urban LU, whereas in Michigan seven of nine individual metrics were significantly correlated; the exceptions being the pFiTol and pFiOmn metrics. When the datasets were combined, all indicators (using both raw and method corrected data) were significantly correlated with both urban and agricultural LUs. Method corrections made no difference in these results (Table 6.4B). However, normalization altered most of the correlations (Table 6.4C and 6.4D), in some cases increasing, in other cases decreasing values, although generally preserving pattern of significance seen the raw data. Visualizations of the correlations for fish and invertebrate taxa richness (Figures 6.7 and 6.8), and for the apparently most sensitive indicators (nFiInt and nEPTFa; Figures 6.9 and 6.10) illustrate the variability in adjustment brought about by the metric normalization.

Because agricultural and urban LUs were themselves significantly correlated ($r = 0.206$, $p < 0.01$), I used GLM ANCOVAs to test regional differences in metric responses to each LU stressor gradient ($p < 0.05$, Appendix 6.5A). When using the raw data for individual indicator metrics, most invertebrate metrics responded significantly to both urban and agricultural LUs (three of four). The exception, as above, being invertebrate family richness (nFaInv), which was correlated with agricultural but not correlated with urban LU. Four of five fish metrics responded significantly to agricultural LU and two of five to urban LU. Region was a significant covariate as a factor (six of nine), or in an interaction term with LU variables in eight of nine cases, indicating that based on the raw sample data metrics there were significant regional differences in stressor-response relationships. When the data were method corrected (Appendix 6.5B), the overall pattern of results was very similar, although effect coefficients were altered in a number of cases. Most metrics again had significant regional differences in their response to LU gradients. These results were also closely similar when large Korean river sites were included in the dataset (Appendix 6.6).

Regional ecological normalization of the data sets controlled explicitly for regional and catchment size covariances (among other variables, Table 6.5 and Appendix 6.7) so it was not surprising that normalized metrics were free of significant regional main effects. However, six of the nine metrics still had significant interaction terms with region (and one or both LU variables) (Table 6.5), indicating statistically significant regional differences in stressor response slope remained as is apparent in the scatter plots (Figures 6.7-6.10 bottom rows). The normalized indicator metrics varied in relative sensitivity to urban and agriculture stresses (Table 6.5B); nFiInt, pFiIns, and all four invertebrate metrics were more sensitive to urban than to agricultural LU while nFiSp, pFiTol, and pFiOmn were more responsive to agriculture. These results were nearly identical when excluded large river sites were restored in the dataset (Appendix 6.7).

LU stressor-response relationships: multi-metrics

Michigan and Korea both use IBI-type (Karr 1981) multi-metrics for their overall assessments with fish data (P51Fi and KIBI, respectively). Both raw and normalized versions of the fish MMIs were negatively correlated with both LU stressors in their respective regions (Table 6.6). The KIBI was more strongly correlated with urban than agricultural LU in S. Korea, while Michigan's P51Fi for fish was more balanced in that respect but had weaker correlations overall compared to the KIBI (Figures 6.11 and 6.12 top rows). When the datasets were combined, scaling differences were conflated with LU response and resulted in erroneously elevated correlation with LU (Table 6.6). Normalization corrected scaling issues (Figures 6.11 and 6.12 middle rows) but suggested differences in underlying stressor-response relationships between S. Korea and Michigan. ANCOVA of the normalized raw fish MMIs indicated that the slopes characterizing the response were significantly higher (i.e. more negative) in Korea than in Michigan for the urban LU gradient (Figure 6.11), and marginally higher (interaction term p= 0.10) for agricultural LU (Figure 6.12, Table 6.7).

The raw invertebrate MMIs (P51Inv and KSI) were also negatively correlated with LU stressor gradients, the relationship being particularly strong in the Korean metric (Table 6.6). The KSI was more strongly correlated with urban than agricultural LU in S. Korea, while Michigan's P51Inv was again more balanced in that respect but had weaker correlations overall compared to the KSI. Directly combining the data again led to scaling issues (top rows of Figures 6.11 and 6.12, Table 6.6), which required normalization to correct. ANCOVA results for the pooled normalized raw MMIs indicated a statistically different response slope in Korea and Michigan although the difference for urban LU was marginal (p= 0.08, Table 6.7).

The regional assessment metrics themselves have been calibrated by their respective users to achieve desired outcomes and therefore reflect different scoring criteria as well as differences in numerical scaling. This, combined with known differences in fish community composition, and potential cultural differences in modes of agricultural and urban land uses (e.g., rice versus corn production techniques, high-rise versus sprawl development patterns) make it difficult to determine whether observed differences in stressor-response relationships are due to differences in sensitivity of the fauna or differences in stressor intensity. The normalization of the indicator metrics allows for convenient algebraic combination (Wiley et al. 2003) so that it is possible to construct normalized multimetrics directly from sets of normalized indicators. Using the same scoring criteria for both regions (standard deviations from modeled expectation) removes scoring differences from the comparison of stressor-response slopes. ANCOVAs of these normalized composite multimetrics (CompFi and nCompInv) indicated significant regional differences remain for the fish community stressor-response relations, but not for the invertebrate community relationships (Table 6.7).

Comparison of ecological condition in Michigan and Korean streams

Ultimately both Michigan and Korean assessment programs aim to provide reasonable estimates of the extent to which local streams and rivers are meeting regional water quality and environmental protection goals as represented by their reference criteria. Normalized regional assessment multimetrics (P51Fi, P51Inv, KIBI, and KSI) and normalized composite multimetrics (CompFi and CompInv) indicated that overall, Korean streams were relatively more degraded compared to Michigan streams (Table 6.8); with lower means and medians in most normalized MMIs except normalized CompInv (Table 6.9). The percentage of impaired sites in S. Korea varied greatly between the fish and invertebrate based metrics (Table 6.8). The normalized KIBI and CompFi all indicated severe impairment of the fish community at many sites (70.1% impaired for normalized KIBI and 66.2% for normalized CompFi). The overall rate of impairment in S. Korea appears to be around 66%, a value not much different from the raw KIBI assessment result (69%).

Rates of impairment varied by stream size (Table 6.10). Overall Korean streams in both smaller and larger regions were more impaired than Michigan streams. When smaller and larger streams were compared with raw MMIs, smaller streams in both regions were generally more impaired than larger streams except KIBI (24.9% for smaller streams and 30.9% for larger streams). However, when raw MMIs were normalized, impairment percentages of all normalized raw MMIs indicated that all larger streams were more likely to be impaired than smaller streams.

Discussion

Several important observations emerge from this comparative case study. First, explicit correction for regional methodological and statistical biases had significant effects on the values of most of the individual and multi-metric indicator variables. As a result, the meaning of the assessment results changed, in some cases dramatically. Second, regional ecological normalization (residualization) and rescaling proved necessary for an unbiased comparison of LU stressor-response relationships across regions. Third, while fish and invertebrate communities were more sensitive to urban LU than to agricultural LU in both regions, stressor-response relationships differed significantly between regions. These and related observations are discussed in more detail below along with their implications for global transferability and comparability of assessment data sets from ecologically distinct regions.

Sampling method biases

All sampling methods vary in respect to efficiency, precision, and bias (ITFM 1995, Merritt and Cummins 1996, Barbour et al. 1999, Cao et al. 2005). Sampling method bias is of concern to both scientists and policy makers who use sample data as a basis for evaluation and management (Barbour et al. 1999, Houston et al. 2002, Wiley et al. 2003, Clarke and Hering 2006, Furse et al. 2006, US EPA 2007). In this case study the biases of regional sampling methods had been already explored (Chapters 4 and 5) and correction factors developed to allow calibration of the samples. The method-corrected data sets used here compensated to some extent for differential sampling biases associated with stream size and other properties. Using this calibrated (methods corrected) data I found that means and taxa counts in both fish and invertebrate richness metrics were notably affected (Appendices 6.2 and 6.3), which indicated that certain metrics showed larger mean differences, whereas others had lower mean differences. However, correlations with stressors were not affected; as mathematically expected since method corrections employed here were linear (count x gear-specifc correction factor). Overall the relative impact of method bias compared to statistical bias appeared in this analysis to be small and explicit correction had only a modest impact on assessment results.

Methodological variations in bias of course become a more critical concern when RBP data from different sources are pooled (ITFM 1995, Barbour et al. 1999, MEA 2005, Furse et al. 2006, US EPA 2007, Riseng et al. 2010, Allan et al. 2013). In my analysis intra-regional correlations with stressors in the pooled data were unaltered but inter-regional differences were magnified although these clearly reflected the different methods employed and not underlying biological responses. The problem has been well understood for some time. For example Barbour et al. 1999 had early-on described various examples of standardization methods to compensate methodological biases, and the Intergovernmental Task Force on Monitoring Water Quality (ITFM) emphasized the importance of data comparability when contrasting data from various RBPs (ITFM 1995).

Attempt to remove methodological bias by gear calibration is rare in the RBP literature although common in fishery assessment studies (see Chapter 4). Recent large-scale RBP studies have generally addressed the problem through method standardization, developing new methods or indicator metrics (ITFM 1995, MEA 2005, Furse et al. 2006, US EPA 2007). In this case, they require new data collections with the new standardized methodologies a solution that is redundant and costly from a public policy perspective.

Correcting for ecological covariance

Statistical corrections had more of an impact than methodological corrections in this regard. Some indicator variables were less affected by bias corrections (i.e. more stable) than others. In particular the KIBI fish multi-metric (An et al. 2005, NIER 2009), and the invertebrate community MBI (Hilsenhoff et al. 1987) were notably more stable in this sense than the other metrics. In addition to the methodological biases, there can also be statistical biases arising from environmental covariances or sample selection bias (Wiley et al. 2003, Cao et al. 2005, Pearl 2009, Hawkins et al. 2010, Schoolmaster et al. 2013). Environmental covariances which influence the spatial patterning of biological communities can be considered to be causal; and can include aspects of hydrogeology, catchment size, river network structure, water temperature, surficial geology, site slope, and land use patterns (Wiley et al. 2003, Zorn et al. 2002, Wehrly et al. 2006). Mathematically, covariance relationships result in complex joint probability distributions which bias correlation and regression coefficients that we normally use to describe the effects of human disturbance on individual metrics. In other words impacts of stressors can be obscured or exaggerated by those environmental covariates (Wiley et al. 2003, Schoolmaster et al. 2013). Thus, interregional data comparability and integrability requires either experimental or statistical control to yield unbiased estimates of response in dependent variables (Pearl 2009, Schoolmaster et al. 2013). Failure to account for natural covariation can lead the wrong interpretations of stressor-response relationships between regions. In this case, for example, raw data analyses led to the conclusion that larger river sites in S. Korea were in relatively better condition than smaller upstream sites. This is a potentially controversial finding since there has been much political debate about river management policies focused on the impoundment of lower river reaches to ease monsoon-related flood damage and improve water quality during dry periods by augmenting natural flow regimes. However, when biases related to catchment size, hydrology, and sampling method were removed, the (normalized) analysis led to the opposite conclusion, i.e., that the down-stream river sites were more impacted by current management than the smaller upstream sites.

Regional ecological normalization (residualization) and rescaling allowed for an unbiased comparison of LU stressor-response relationships from both regions. The degree of similarity in stressor-response covaried with taxonomic (evolutionary) similarity of the regional communities. For the invertebrate community metrics there was little or no difference in response to either urban or agricultural land use gradients between the two regions; neither was there much difference at order or family level in community composition between S. Korea and Michigan. In contrast, the S. Korean fish community and the Michigan fish community composition were fundamentally different (species overlap; 4.2% of Michigan species and 3.8% of Korean species), and their responses to both urbanization and agricultural land use were quite different. MLR models (Appendices 6.8 though 6.11) were used to adjust expectations for statistically significant covariates for each metric. All of the models (58 models in total) included significant terms for site drainage area. Site slope, water temperature, and stream flow yields were also often important variables. Specifically, low flow yield (Q90Y) was important in 10 out of 18 models for Michigan, whereas high flow yield (Q10Y) was important in 12 out of 18 models for S. Korea. The contrast is striking, and suggests that the two regions may have fundamentally different hydrologic constraints (monsoon flows, groundwater supported base flows) shaping their biological communities. While it is not possible to know whether all such statistical biases are accounted for in the normalization modeling, recent simulations studies have examined several scenarios using both reference-set residualization (Whittier et al. 2007, Stoddard et al. 2008), and whole-set residualization (called “regional normalization” by Wiley et al. 2003).

Results of these controlled simulations indicated that regional normalization not only produced more accurate, precise, and efficient adjustments to the specified covariates, but also eliminated the need for classification of the disturbance state of sites into "reference" and "impacted" sites (Schoolmaster et al 2013).

For the sake of consistency, I have mostly presented here analyses of normalized (residualized) data sets that had already been method-corrected. However, all analyses were run on both normalized raw and normalized method-corrected data, and these showed almost identical correlation (Table 6.4) and ANCOVA results (Table 6.5) in relation to LU stressors. This suggests that regional normalization (residualization) may not require prior data correction for the methodological biases from disparate sampling protocols. This should not be surprising since the data are re-centered and re-scaled in the process, removing any systematic bias between the regions (Wiley et al 2003, Riseng et al. 2006).

LU stressor-response relationships

Normalized regional MMIs indicated that both fish and invertebrate communities were more sensitive to urban LU gradients than agricultural LU, and this was true in both regions (Table 6.7). Also, response slopes of normalized fish and invertebrate MMIs were significantly higher in Korean than in Michigan for both urban and agricultural LU gradients. This did not surprise me given the background of human population density and cultural practices with respect to LU management in both regions. Agricultural land use is a well known anthropogenic stressors (Allan and Johnson 1997, MEA 2005, Riseng et al. 2010, US EPA 2006). However, the same proportion of agricultural land use in a catchment does not cause the same effect on aquatic ecosystems in different regions (Riseng et al. 2011). For an example, a densely populated urban region may have more intensified uses of pesticides and higher sewage-related nutrient exports than an area with similar proportions of urban LU but with lower population densities. LU patterns, LU intensities, and LU-related technologies are all culturally mediated, and this is especially so in the cases of agriculture where resources and human market preferences cast a long shadow. Because of this, stressor-response gradients in land use need not necessarily be similar in different regions.

On the other hand, a different interpretation can be made based on the results of normalized composite multimetrics (Figures 6.9 and 6.10, Table 6.7). The responses of the composite multimetrics (CompFi and CompInv) are free of any regionally imposed interpretive criteria (unlike the normalized regional MMIs) and they suggest that the degree of similarity in stressor-response relationships covaried with taxonomic (evolutionary) similarity of the regional communities. Invertebrate communities in Michigan and S. Korea showed little or no difference in response to either urban or agricultural LU gradients, whereas fish communities showed significantly different responses. Likewise, the evolutionary background of invertebrates in Michigan and Koran were quite similar (~76% overlap in families, 80 % in orders); but the fish communities were very dissimilar (42% and 50%, respectively). In general, response slopes (indicative of sensitivity) in Korea were higher than in Michigan, except for normalized CompInv MMI response to urban LU gradients. Taken together these two results suggest that both differences in LU intensity and in biogeographic history in two regions may contribute to observed differences in the stressor-response relationships.

Implications for global method transferability and inter-regional assessment

This case study provided a useful context for exploring issues related to inter-regional ecological assessment data. Datasets from geographically and ecologically disparate regions carry with them various types of methodological and statistical biases. The datasets from

Michigan and S. Korea were typical in this regard. Both methodological biases associated with field sampling protocols, and statistical biases (principally covariances with catchment size and hydrology) were encountered in these data. Furthermore, there were regional differences in metric scaling, and narrative criteria for both the fish and invertebrate MMIs. Without some way to control for these types of differences, ecosystem assessments need to be conducted within the boundary of geomorphologically and ecologically comparable (homogenous?) study units using standardized protocols and metrics. Thus, many existing larger-scale (e.g., US EPA 2007), international (e.g., Furse et al. 2006), or global ecosystem (e.g., MEA 2005) assessments can provide only the relative impairment of sites and stressor-response relationships within and between more homogeneous sub-regions. In contrast, this comparative study illustrated an approach for eliminating most statistical and methodological biases, thus allowing for a more direct comparison of ecological impairment and LU stressor-response relationships between two international regions.

The assessment protocols and metrics used in S. Korea are an excellent example of issues that can arise in EATT transfer between regions. The KIBI is based on the North American IBI for fish, but has been extensively re-calibrated and some of the constituent metrics replaced to reflect the structure of fish communities in S. Korea (An et al. 2005, NIER 2009). The invertebrate multimetric (the KSI) is based on the European saprobic valency concept (Zelinka and Marvan 1961), again with local adaptations; although in this case the general similarities of the European and Korean, (and North American) aquatic insect fauna are high. Of the four regional MMIs discussed here, the KIBI raw metric was most highly correlated with the average of the normalized fish metrics suggesting it was already well calibrated to the local fish community response and relatively free of covariances with size, hydrology, or other factors. In contrast the invertebrate KSI metric was least correlated with its normalized counterparts suggesting significantly more bias in this metric (r= .85 and .75 for KIBI and KSI respectively).

Perhaps the most interesting implication of these results, relevant to both matters of EATT transfer and data comparability, is the relationship observed between similarity of response to LU stress gradients, and the taxonomic similarity of the indicator communities. If the response of biological communities to LU change varies (non-trivially) with composition of the indicator community then indicator choices, community susceptibility to stresses, and the validity of assessment criteria might all be expected to have geography and scaling that reflects zoogeographic and recent evolutionary history. This would imply a kind of spatial bounding in the scale at which stressor-response relationships should be stable, and that in turn would have both assessment and regulatory consequences. It is just this kind of comparative analysis, using data from very different parts of the world, which would be necessary to evaluate such a hypothesis. And of course to carry out such analyses we need to control for methodological and statistical biases.

Summary

In this chapter I have examined the comparability of two datasets obtained from geographically distinct regions (S. Asia and N. America) and produced using two different rapid assessment protocols. I concluded that ecological data from different geographical regions are not directly comparable. However a regional normalization approach (Wiley et al. 2003) was a useful tool to correct methodological and statistical biases and standardize outputs. Ecological assessment of two regions using normalized data indicated that S. Korean streams were overall more seriously degraded than Michigan streams. Also, normalized scores for certain indicator metrics showed that their LU stressor-response relationships were significantly different so that interregional/international regulatory policies and management in larger scales should be cautious depending on their goals and targets for watershed management and resource conservation.

Overall, this comparative international case study of the transferability and comparability of EATTs demonstrates the degree to which regional methodologies, and differences in physiography and hydrology, can skew and obscure the meaning of ecological assessment data. Specifically, recognizing the role of potential biases in assessment will enable policy makers and researchers to compensate for the inherent limitations related to site geomorphology, biology, and methodologies (e.g., MEA 2005, Furse et al. 2006, US EPA 2007). In fact, these constraints have not been correctly recognized in many ecological studies and are conflated with true anthropogenic impacts, resulting in inappropriate analysis and conclusions. Thus, comparability and transferability in ecosystem-assessment techniques and tools should be examined in every study that uses regional-scale data or integrated data sets from different regions/methods/or indicators. In this regard, the techniques used here can hopefully guide by way of example the path towards more accurate assessment analyses.

Table 6.1. Definition of acronyms frequently used in this study.

Category	Acronym	Full description or definition
General	RBP	Rapid bioassessment program
	MDEQP51	Michigan Department of Environmental Quality Procedure 51
	KNAEMP	The Korean Nationwide Aquatic Ecological Monitoring Program
	MLR	Multiple linear regression
Landscapes	DA	Drainage area (km^2)
	LU	Land use
	xUrb	Percentage of urban land use in catchment (%)
	xAg	Percentage of Agricultural land use in catchment (%)
	Q10Y	High (10% exceedance frequency) stream flow yield ($m^3/sec/km^2$)
	Q50Y	Median (50% exceedance frequency) stream flow yield ($m^3/sec/km^2$)
	Q90Y	Low (90% frequency) stream flow yield ($m^3/sec/km^2$)
Individual metrics	nFiSp	Number of total fish species
	nFiInt	Number of intolerant fish species
	pFiTol	Percentage of tolerant fish individuals
	pFiOmn	Percentage of omnivorous fish individuals
	pFiIns	Percentage of insectivorous fish individuals
	nFaInv	Number of total invertebrate families
	EPT	Invertebrate groups of Ephemeroptera, Plecoptera, and Trichoptera
	nEPTFa	Number of EPT families
	pEPTIn	Percentage of total individuals that were EPT
	MBI	MBI Biotic Index score
Multimetric indicies	MMI	Multimetric Index
	KIBI	Korean Index of Biological Integrity for fish
	KSI	Korean Saprobic Index for benthic macroinvertebrates
	P51Fi	MDEQP51 for fish
	P51Inv	MDEQP51 for benthic macroinvertebrates
	CompFi	Overall composite fish MMI
	CompInv	Overall composite invertebrate MMI

Table 6.2. List of biological indicator metrics (fish and benthic macroinvertebrates) used for this study and their equations employed to each metric to correct numerical differences by sampling gear or methods between Michigan and S. Korean regions. MMI and MKO indicate individual metrics of Michigan and S. Korea, respectively. cMMI and cMKO indicate method corrected individual metrics of each region.

Biological groups	Indicator metrics	Equations
Fish	Number of total fish species (nFiSp)	$cM_{KO} = 1.2429 * M_{KO}$
	Number of intolerant fish species (nFiInt)	$cM_{KO} = 1.2678 * M_{KO}$
	Percentage of tolerant fish individuals (pFiTol)	$cM_{KO} = 0.8967 * M_{KO}$
	Percentage of omnivorous fish individuals (pFiOmn)	$cM_{KO} = 0.9779 * M_{KO}$
	Percentage of insectivorous fish individuals (pFiIns)	$cM_{KO} = 0.8871 * M_{KO}$
	Korean Index of Biological Integrity (KIBI) score	
	MDEQP51 for fish (P51Fi) score	
Benthic macroinvertebrates	Number of total invertebrate families (nFaInv)	$cM_{MI} = 1.3002 * M_{MI}$
	Number of EPT families (nEPTFa)	$cM_{MI} = 1.2952 * M_{MI}$
	Percentage of total individuals that were EPT (pEPTIn)	$cM_{MI} = 0.9760 * M_{MI}$
	MBI Biotic Index (MBI) score	$cM_{MI} = 0.9366 * M_{MI}$
	Korean Saprobic Index (KSI) score	
	MDEQP51 (P51Inv) score	

Table 6.3. Summary statistics of landscape variables between Michigan and South Korean regions used for the study. SD indicates standard deviation.

Variables	n	Mean	Median	SD	Min	Max
Michigan region						
Drainage Area; km^2	803	173.7	55.9	330.8	0.5	3334.4
Average wetted width; m	802	24.3	16	24.4	1.2	200.0
Average depth; m	801	1.15	1.00	0.83	0.10	10.00
Water temperature; ºC	758	17.6	17.8	3.9	0.6	28.3
Percent of urban land use (xUrb)	803	6.08	4.27	6.48	0.00	52.25
Percent of agricultural land use (xAg)	803	39.32	41.01	26.39	0.00	94.33
High-flow yield (Q10Y)	803	0.021	0.020	0.005	0.003	0.037
Median-flow yield (Q50Y)	803	0.007	0.007	0.003	0.000	0.018
Low-flow yield (Q90Y)	803	0.003	0.002	0.002	0.000	0.012
Korean region (large river sites excluded)						
Drainage Area; km^2	633	337.0	114.7	566.5	0.6	3469.3
Average wetted width; m	633	46.5	25.0	65.5	0.2	500.0
Average depth; m	633	0.47	0.25	0.66	0.00	5.00
Water temperature; ºC	633	15.6	16.0	4.5	1.0	27.6
Percent of urban land use (xUrb)	633	8.75	3.63	14.05	0.00	100.00
Percent of agricultural land use (xAg)	633	22.03	18.42	16.62	0.00	83.10
High-flow yield (Q10Y)	633	0.084	0.074	0.049	0.012	0.447
Median-flow yield (Q50Y)	633	0.028	0.022	0.023	0.004	0.297
Low-flow yield (Q90Y)	633	0.015	0.011	0.013	0.002	0.094

Table 6.4. Pearson correlation coefficients between biological indicator metrics and LU stressors for Michigan, Korean (large river sites excluded), and combined (both regions) datasets. Bold indicates significance at p ≤ 0.05 and bold and italics indicate significance at p ≤ 0.01.

Indicator Metrics	Michigan region			S. Korean region			Combined data		
	xUrb	xAg	ln(DA)	xUrb	xAg	ln(DA)	xUrb	xAg	ln(DA)
A. Raw data									
nFiSp	***-0.174***	***0.218***	***0.499***	***-0.19***	0.063	***0.386***	***-0.182***	***0.205***	***0.397***
nFiInt	***-0.152***	***-0.28***	***0.334***	***-0.352***	***-0.372***	***0.169***	***-0.289***	***-0.250***	***0.219***
pFiTol	0.014	***0.462***	-0.047	***0.392***	***0.419***	***0.111***	***0.274***	***0.336***	***0.069***
pFiOmn	0.022	***0.266***	***-0.222***	***0.324***	***0.396***	**0.086**	***0.247***	***0.153***	0.020
pFiIns	**-0.081**	-0.049	***0.339***	***-0.313***	***-0.456***	***-0.192***	***-0.234***	***-0.201***	0.052
nFaInv	***-0.286***	***-0.221***	***0.352***	***-0.337***	***-0.196***	**0.083**	***-0.295***	***0.072***	**0.068**
nEPTFa	***-0.278***	***-0.452***	***0.265***	***-0.343***	***-0.344***	**0.093**	***-0.320***	***-0.285***	***0.134***
pEPTIn	***-0.220***	***-0.377***	***0.169***	***-0.243***	***-0.354***	***0.109***	***-0.228***	***-0.321***	***0.136***
MBI	***0.211***	***0.494***	-0.043	***0.326***	***0.227***	***-0.104***	***0.303***	***0.267***	***-0.069***
B. Method corrected data									
nFiSp	***-0.174***	***0.218***	***0.499***	***-0.19***	0.063	***0.386***	***-0.167***	***0.122***	***0.444***
nFiInt	***-0.152***	***-0.28***	***0.334***	***-0.352***	***-0.372***	***0.169***	***-0.289***	***-0.289***	***0.237***
pFiTol	0.014	***0.462***	-0.047	***0.392***	***0.419***	***0.111***	***0.255***	***0.372***	0.049
pFiOmn	0.022	***0.266***	***-0.222***	***0.324***	***0.396***	**0.086**	***0.245***	***0.159***	0.016
pFiIns	***-0.081***	-0.049	***0.339***	***-0.313***	***-0.456***	***-0.192***	***-0.235***	***-0.162***	0.047
nFaInv	***-0.286***	***-0.221***	***0.352***	***-0.337***	***-0.196***	**0.083**	***-0.266***	***0.128***	0.033
nEPTFa	***-0.278***	***-0.452***	***0.265***	***-0.343***	***-0.344***	**0.093**	***-0.306***	***-0.220***	***0.104***
pEPTIn	***-0.220***	***-0.377***	***0.169***	***-0.243***	***-0.354***	***0.109***	***-0.226***	***-0.325***	***0.139***
MBI	***0.211***	***0.494***	-0.043	***0.326***	***0.227***	***-0.104***	***0.316***	***0.202***	-0.042
C. Normalized raw data									
nFiSp	***-0.148***	***0.108***	***0.126***	***-0.146***	-0.073	***-0.134***	***-0.161***	***0.126***	**-0.058**
nFiInt	***-0.129***	***-0.113***	0.009	***-0.463***	***-0.462***	***-0.183***	***-0.348***	0.030	***-0.214***
pFiTol	-0.027	***0.383***	0.042	***0.352***	***0.372***	***0.122***	***0.242***	***0.154***	***0.171***
pFiOmn	-0.008	***0.205***	0.020	***0.236***	***0.332***	0.077	***0.179***	**0.055**	***0.142***
pFiIns	***-0.136***	-0.047	0.007	***-0.429***	***-0.403***	***-0.153***	***-0.333***	**0.058**	***-0.189***
nFaInv	***-0.265***	***-0.218***	0.005	***-0.357***	***-0.184***	-0.025	***-0.262***	***-0.270***	0.036
nEPTFa	***-0.304***	***-0.241***	0.005	***-0.352***	***-0.363***	***-0.132***	***-0.312***	***-0.269***	**-0.055**
pEPTIn	***-0.239***	***-0.171***	0.008	***-0.295***	***-0.405***	***-0.139***	***-0.280***	***-0.153***	***-0.101***
MBI	***0.265***	***0.278***	0.006	***0.321***	***0.273***	0.078	***0.283***	***0.254***	0.039
D. Normalized method corrected data									
nFiSp	***-0.148***	***0.108***	***0.126***	***-0.146***	-0.072	***-0.134***	***-0.161***	***0.126***	**-0.058**
nFiInt	***-0.129***	***-0.113***	0.009	***-0.349***	***-0.5***	***-0.244***	***-0.286***	0.028	***-0.245***
pFiTol	-0.027	***0.383***	0.042	***0.366***	***0.389***	***0.127***	***0.252***	***0.141***	***0.180***
pFiOmn	-0.008	***0.205***	0.020	***0.236***	***0.332***	0.077	0.179	**0.055**	***0.142***
pFiIns	***-0.136***	-0.047	0.007	***-0.429***	***-0.405***	***-0.153***	***-0.333***	**0.058**	***-0.189***
nFaInv	***-0.265***	***-0.217***	0.005	***-0.357***	***-0.184***	-0.025	-0.263	***-0.269***	0.036
nEPTFa	***-0.303***	***-0.232***	0.006	***-0.352***	***-0.363***	***-0.132***	***-0.314***	***-0.259***	**-0.058**
pEPTIn	***-0.240***	***-0.171***	0.007	***-0.295***	***-0.405***	***-0.139***	***-0.280***	***-0.154***	***0.100***
MBI	***0.265***	***0.275***	***0.005***	***0.321***	***0.273***	0.078	0.286	***0.243***	0.043

Table 6.5. Coefficients and their statistical significance from GLM ANCOVA tests of normalized scores of biological indicator metrics with urban and agricultural LUs between Michigan and Korean (large river site excluded) datasets. Region (Michigan and Korea) was used as fixed factor and LU stressors (xUrb and xAg) were used as covariates. Bold indicates significance at $p \leq 0.05$ and bold and italics indicate significance at $p \leq 0.01$. Coefficients for Region, Region*ln(xUrb), and Region*ln(xAg) were summarized for Korean data and coefficients of these for Michigan data could be calculated by multiplying by -1 to coefficients of Korean data.

Covariates	Biological indicator metrics								
	ln(nFiSp)	ln(nFiInt)	ln(pFiTol)	ln(pFiOmn)	ln(pFiIns)	ln(nFaInv)	ln(nEPTFa)	ln(pEPTIn)	ln(MBI)
A. *Normalized raw data (Michigan versus Korea (large river sites excluded))*									
Region	-0.1364	-0.06396	-0.01325	-0.04288	0.03246	0.03289	-0.005958	-0.005408	-0.0003017
ln(xUrb)	-0.06134	***-0.2885***	***0.2611***	***0.2428***	***-0.2814***	***-0.3825***	***-0.3769***	***-0.2966***	***0.391***
Region*ln(xUrb)	0.04145	***-0.2714***	***0.3184***	***0.1864***	***-0.1966***	0.03616	0.0329	-0.007548	0.008359
ln(xAg)	***0.2022***	-0.08661	***0.4123***	***0.33***	-0.05677	-0.06054	**-0.111**	***-0.1209***	***0.1886***
Region*ln(xAg)	**-0.06693**	***-0.1509***	0.03494	***0.1***	***-0.1586***	0.04593	-0.02593	***-0.09166***	0.01266
ln(xUrb)*ln(xAg)	***-0.07076***	***-0.06942***	-0.03752	-0.04703	-0.04305	-0.007345	-0.03527	-0.03115	-0.0009251
B. *Normalized method corrected data (Michigan versus Korea (large river sites excluded))*									
Region	-0.1366	0.003458	-0.01402	-0.04286	0.03235	0.03291	-0.006142	-0.00541	0.02636
ln(xUrb)	-0.06212	***-0.2722***	***0.2802***	***0.2428***	***-0.2819***	***-0.3828***	***-0.3736***	***-0.2967***	***0.3909***
Region*ln(xUrb)	0.04154	***-0.238***	***0.3324***	***0.1867***	***-0.1974***	0.03625	0.03079	-0.007388	0.008057
ln(xAg)	***0.2017***	***-0.168***	***0.4276***	***0.3301***	-0.05811	-0.06033	**-0.1062**	***-0.1211***	***0.187***
Region*ln(xAg)	**-0.06683**	***-0.2205***	0.04693	***0.1002***	***-0.1601***	0.0456	-0.02981	***-0.09148***	0.01394
ln(xUrb)*ln(xAg)	***-0.07038***	**-0.06185**	-0.03969	-0.04695	-0.04321	-0.007273	-0.03581	-0.03116	-0.0007508

Table 6.6. Pearson correlation coefficients between assessment metrics (raw and normalized) and LU stressors for Michigan, Korean, and combined (both regions) datasets. Large river sites were removed for Korean data. Bold indicates significance at $p \leq 0.05$ and bold and italics indicate significance at $p \leq 0.01$.

Indicator Metrics	Michigan region			S. Korean region			Combined data		
	xUrb	xAg	ln(DA)	xUrb	xAg	ln(DA)	xUrb	xAg	ln(DA)
Raw fish MMI	**-0.101**	***-0.197***	***0.188***	***-0.389***	***-0.377***	**-0.095**	**-0.061**	***-0.604***	0.051
Normalized raw fish MMI	-0.077	***-0.180***	0.003	***-0.443***	***-0.358***	***-0.149***	***-0.383***	-0.027	***-0.114***
Raw invertebrate MMI	***-0.221***	***-0.287***	***0.218***	***-0.373***	***-0.298***	***0.109***	***-0.143***	***-0.376***	***0.229***
Normalized raw invertebrate MMI	***-0.212***	***-0.150***	0.003	***-0.344***	***-0.332***	***-0.146***	***-0.304***	***-0.104***	***-0.113***
Normalized CompFi Score	***-0.157***	***0.197***	0.002	***-0.453***	***-0.452***	***-0.149***	***-0.356***	-0.002	***-0.205***
Normalized CompInv Score	***-0.322***	***-0.264***	0.002	***-0.380***	***-0.346***	-0.074	***-0.339***	***-0.274***	-0.031

Table 6.7. Coefficients, estimated slopes, and slope difference calculated from GLM ANCOVA tests of normalized assessment scores with urban and agricultural LUs between Michigan and Korean datasets. Region (Michigan and Korea) was used as fixed factor and LU stressors (xUrb and xAg) were used as covariates. Estimated slopes of were estimated from coefficient and interaction coefficient with region. Slope difference was calculated by estimated slope of Korea divided by estimated slope of Michigan. Statistical significance cut-off was 0.1. Bold coefficient values indicate significance at p ≤ 0.10 and bold and italic values indicate significance at p ≤ 0.05.

MMIs	LU stressors	Coefficient	Interaction Coefficient with region	Estimated slope		Regional difference	Slope difference	R2	F-ratio
				SubKO	AllMI				
Normalized fish MMI	xUrb	***-0.2968***	***0.2023***	-0.4991	-0.0945	Yes	5.28148	32.2	85.068
	xAg	***-0.1215***	**0.0552**	-0.1767	-0.0663	Maybe	2.66516		
Normalized invert MMI	xUrb	***-0.3515***	**0.05821**	-0.40971	-0.29329	Maybe	1.39695	21.2	61.447
	xAg	***-0.1344***	***0.07928***	-0.21368	-0.05512	Yes	3.87663		
Normalized CompFi	xUrb	***-0.2232***	***0.1835***	-0.4067	-0.0397	Yes	10.2445	56.4	282.349
	xAg	***-0.09988***	***0.1066***	-0.20648	0.00672	Yes	-30.72619		
Normalized CompInv	xUrb	***-0.3595***	0.01645	-0.34305	-0.37595	No	0.91249	19.9	56.405
	xAg	***-0.1098***	0.01757	-0.12737	-0.09223	No	1.38100		

Table 6.8. Percentages of sites for each stream health class for normalized raw and composite multimetric indices.

	Normalized raw fish MMIs		Normalized raw invertebrate MMIs		Normalized CompFi		Normalized CompInv	
	P51Fi (n= 449)	KIBI (n= 633)	P51Inv (n= 757)	KSI (n= 619)	MI (n= 707)	KO (n= 612)	MI (n= 744)	KO (n= 624)
Exellent	10.9	4.9	9.5	0.2	5.2	0.0	0.5	0.6
Good	22.9	16.0	30.3	26.2	57.9	12.1	25.8	32.7
Fair	18.3	9.0	16.9	22.6	22.5	21.7	26.7	19.9
Poor	36.3	28.4	28.4	20.8	13.0	33.0	31.6	26.8
Very Poor	11.6	41.7	14.9	30.2	1.4	33.2	15.3	20.0
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 6.9. Summary statistics of normalized-raw and -composite multimetrics for Michigan and Korea regions. SD indicates standard deviation.

Multimetric indices	n	Mean	Median	SD	Min	Max
Michigan region						
Normalized P51Fi	449	-0.8645	-0.9271	1.0114	-3.1886	2.0387
Normalized P51Inv	757	-0.8509	-0.8386	1.0320	-3.6357	1.7078
Normalized ComFi	707	-0.3617	-0.2949	0.6253	-2.8763	1.2836
Normalized ComInv	744	-1.1029	-0.9467	0.8902	-4.9674	1.1312
Korean region						
Normalized KIBI	633	-1.6531	-1.7125	1.2422	-4.4000	1.7218
Normalized KSI	619	-1.4001	-1.0540	1.1615	-4.1572	0.7554
Normalized ComFi	612	-1.5439	-1.3710	0.9216	-3.6568	0.4279
Normalized ComInv	624	-1.0621	-0.8629	0.9610	-3.5859	0.6581

Table 6.10. Percentages of impaired streams sites for small and large streams based on stream health classification with raw and normalized assessment scores. Impairment classification of raw assessment scores for fish and benthic macroinvertebrates were based on P51Fi and P51Inv classes for Michigan region and KIBI and KSI classes for S. Korean region. These regional data were regionally normalized and reclassified for the impairment status based on normalized assessment scores. Large stream sites were arbitrarily defined as having catchment areas bigger than 500 km^2. SM and LR stand for smaller and larger streams, respectively.

	Percentage of stream impairment for fish				Percentage of stream impairment for invertebrates			
	Raw MMI		Normalized raw MMI		Raw MMI		Normalized raw MMI	
	Smaller streams	Larger streams	Smaller streams	Larger streams	Smaller streams	Larger streams	Smaller streams	Larger streams
Michigan	23.5	6.5	7.2	10.9	10.5	5.7	46.5	50.7
Korea	24.9	30.9	24.1	29.1	22.4	13.6	46.0	50.5

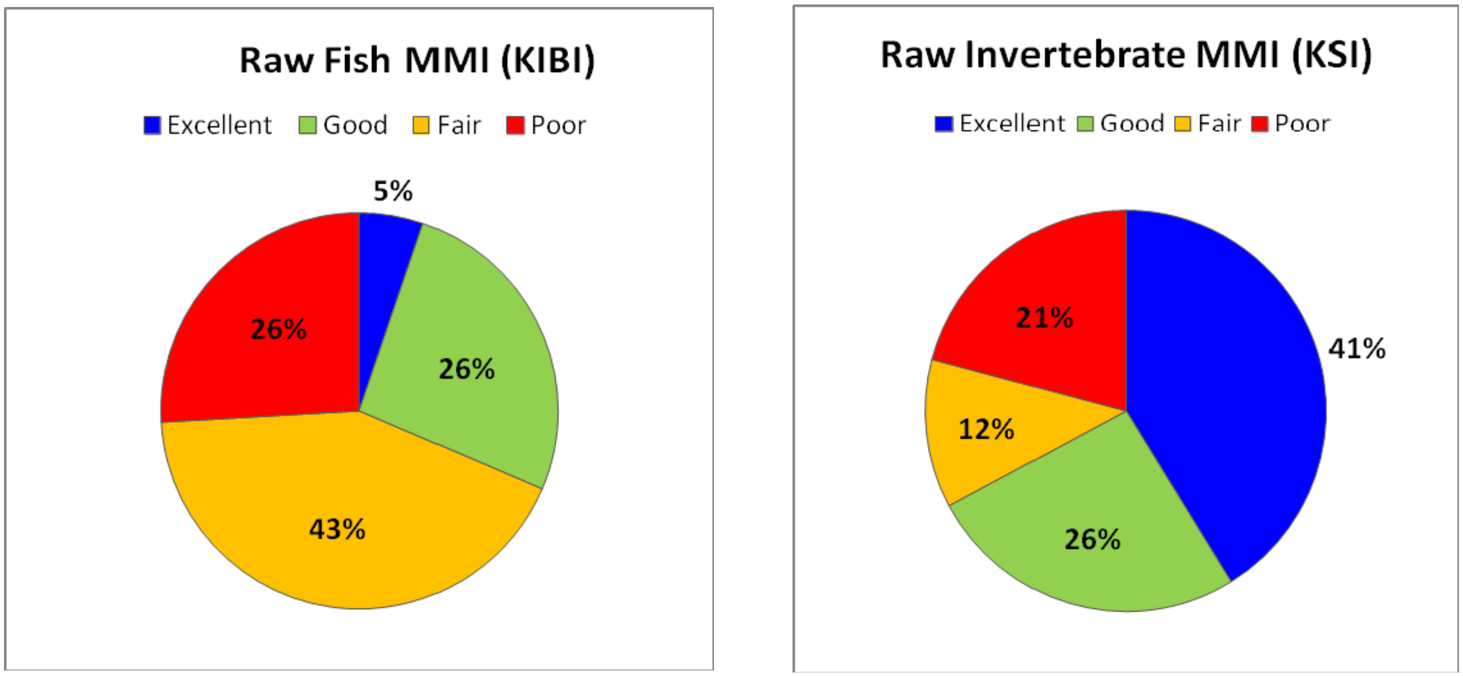


Figure 6.1. Pie charts of stream health assessment classification of Korean streams (n= 633) produced by each regional rapid bioassessment protocols for fish and benthic macroinvertebrates. All assessment data were obtained from the Korean National Aquatic Ecological Monitoring Program (NIER 2009).

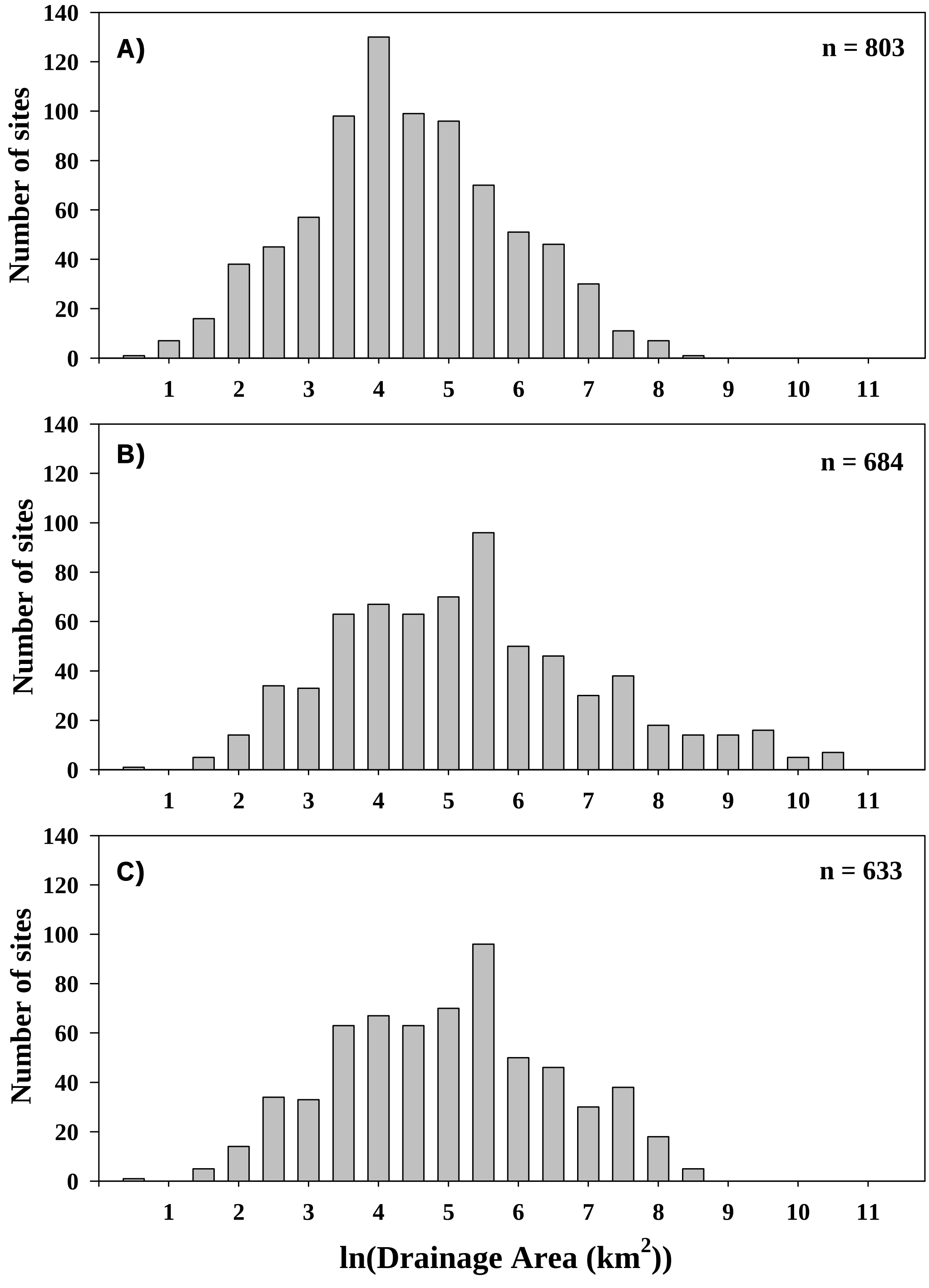


Figure 6.2. Comparison of the frequency distribution of site drainage-area (ln(drainage area+1)) for each region (A: Michigan region (803 sites), B: Korean region with large river sites included (684 sites), and C: Korean region with large river sites excluded (633 sites).

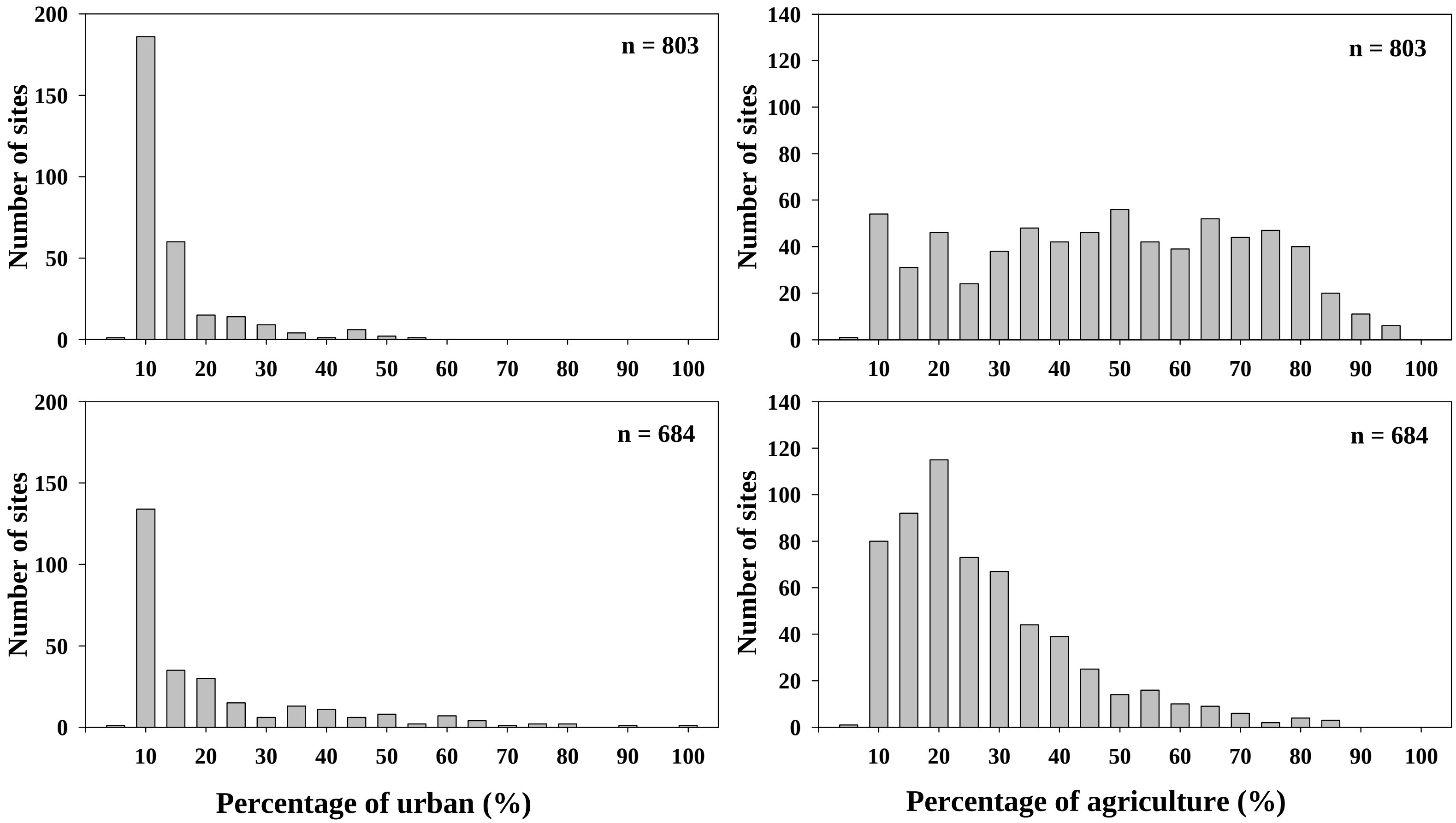


Figure 6.3. Comparison of the urban and agricultural LU frequency distribution between Michigan (n= 803) and Korean (n=684) regions.

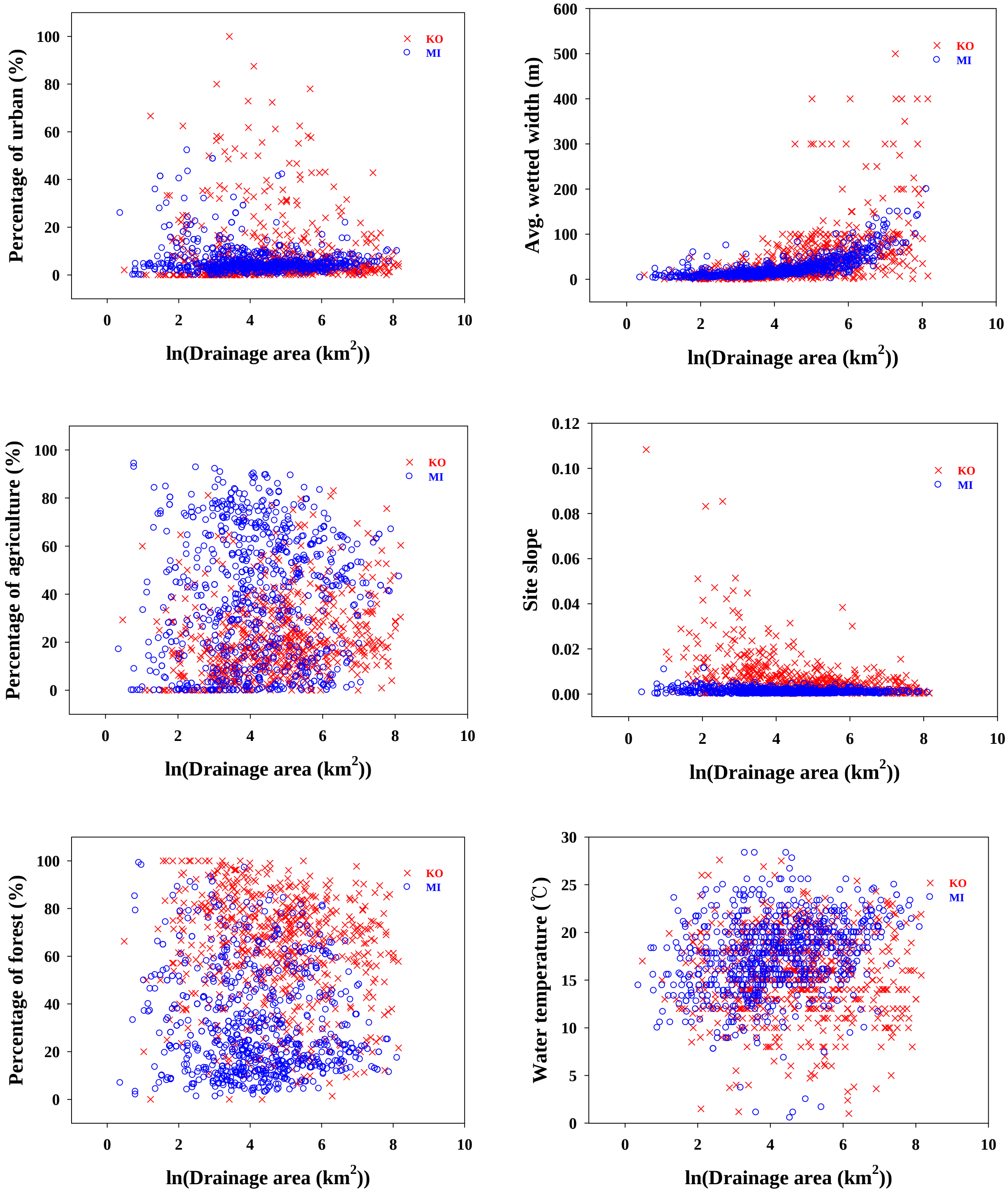


Figure 6.4. Comparison of the distribution of landscape variables against site drainage area between Michigan and Korean regions.

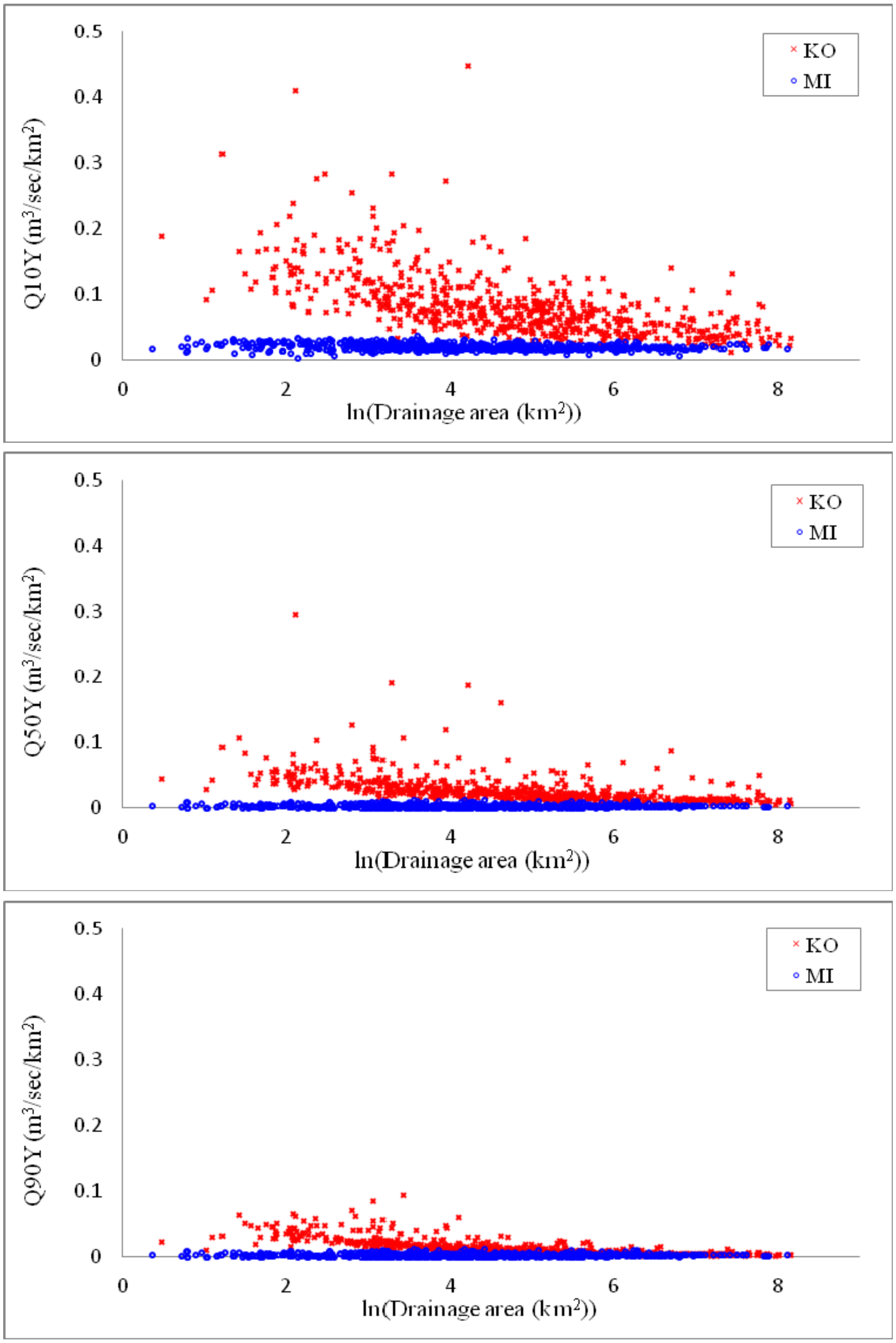


Figure 6.5. Comparision of high (10%), median (50%), and low (90%) frequency flow yields ($m^3/km^2/sec$) between Michigan and Korean regions.

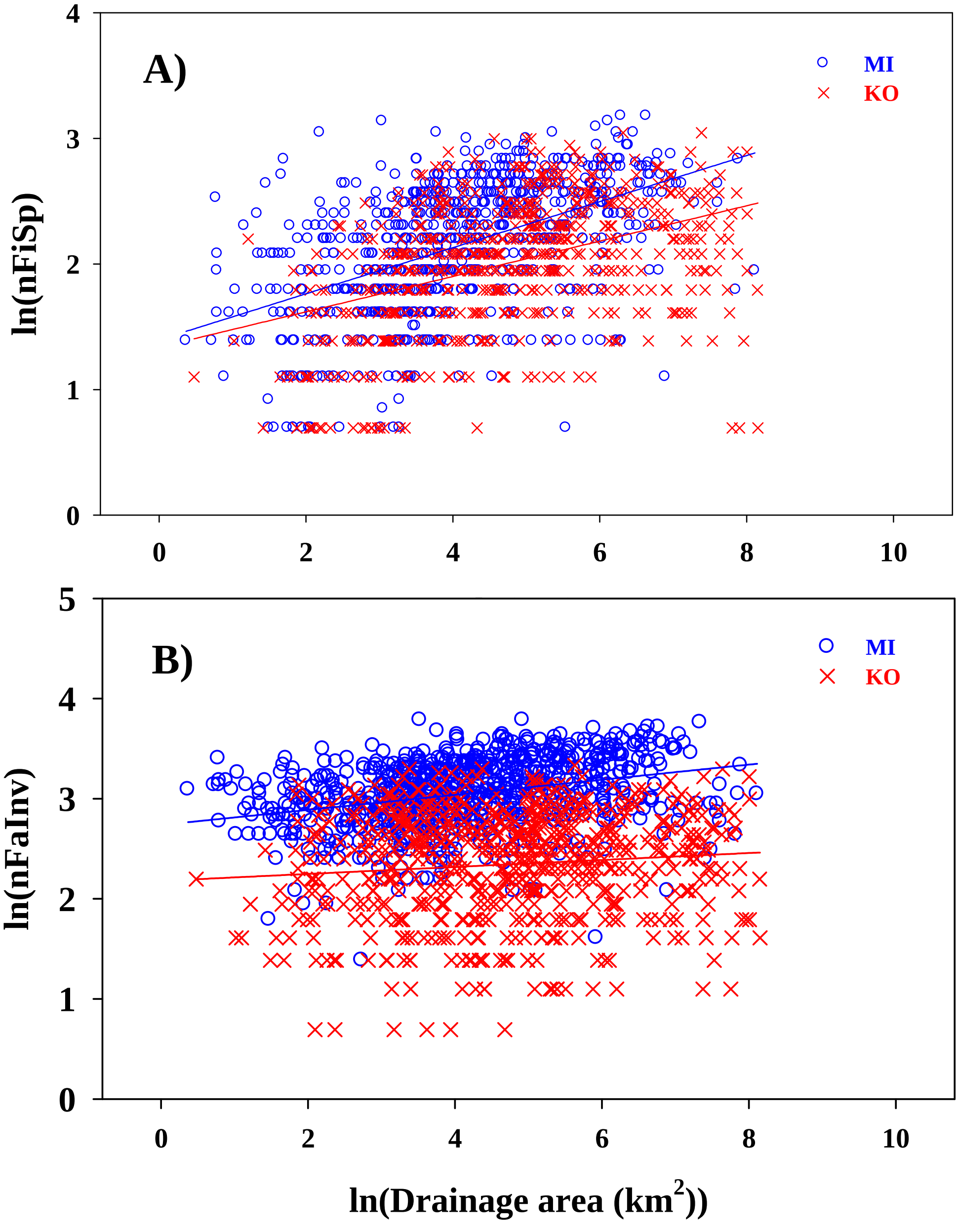


Figure 6.6. Comparison of Michigan (blue circles) and Korean (red x marks) datasets. Axes are A) number of total fish species (nFiSp) and B) number of total invertebrate families (nFaInv) plotted against natural log of site drainage area. All fish and invertebrate data were from raw data collected by regional sampling methods. Large river sites for Korean dataset were excluded.

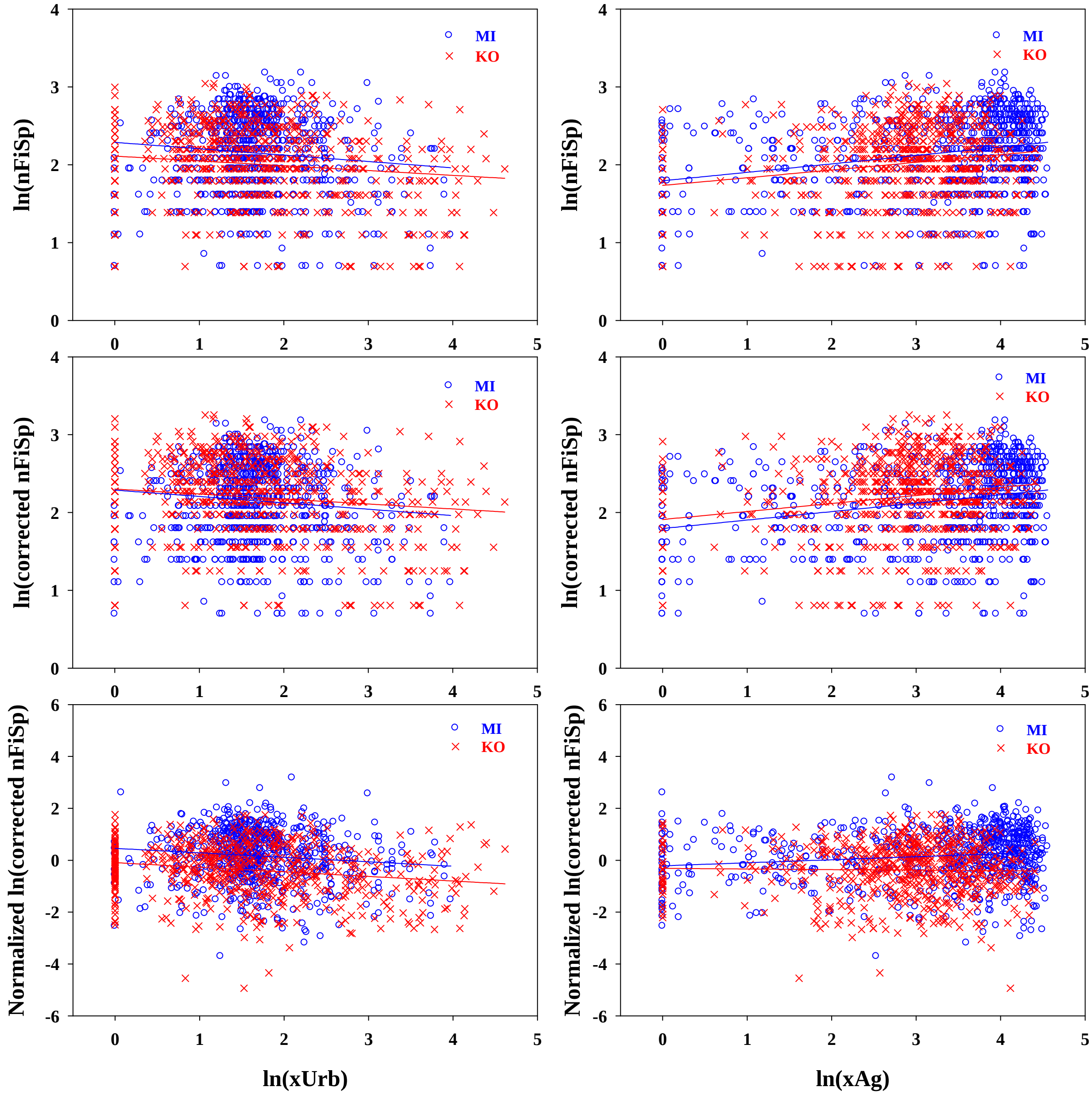


Figure 6.7. Scatter plots of number of total fish species (nFiSp) against urban and agricultural LUs between Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data).

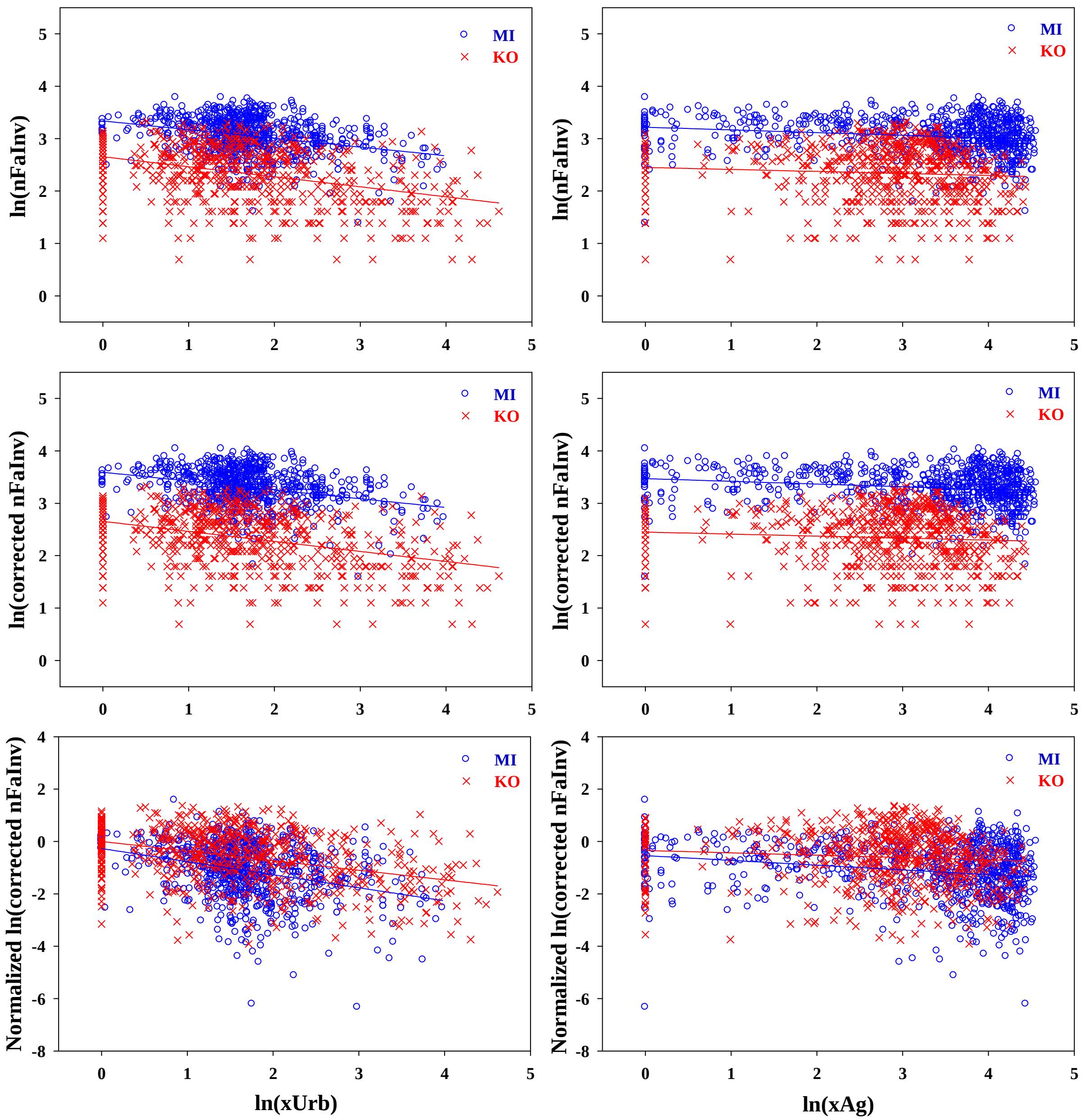


Figure 6.8. Scatter plots of number of total invertebrate families (nFaInv) against urban and agricultural LUs Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data).

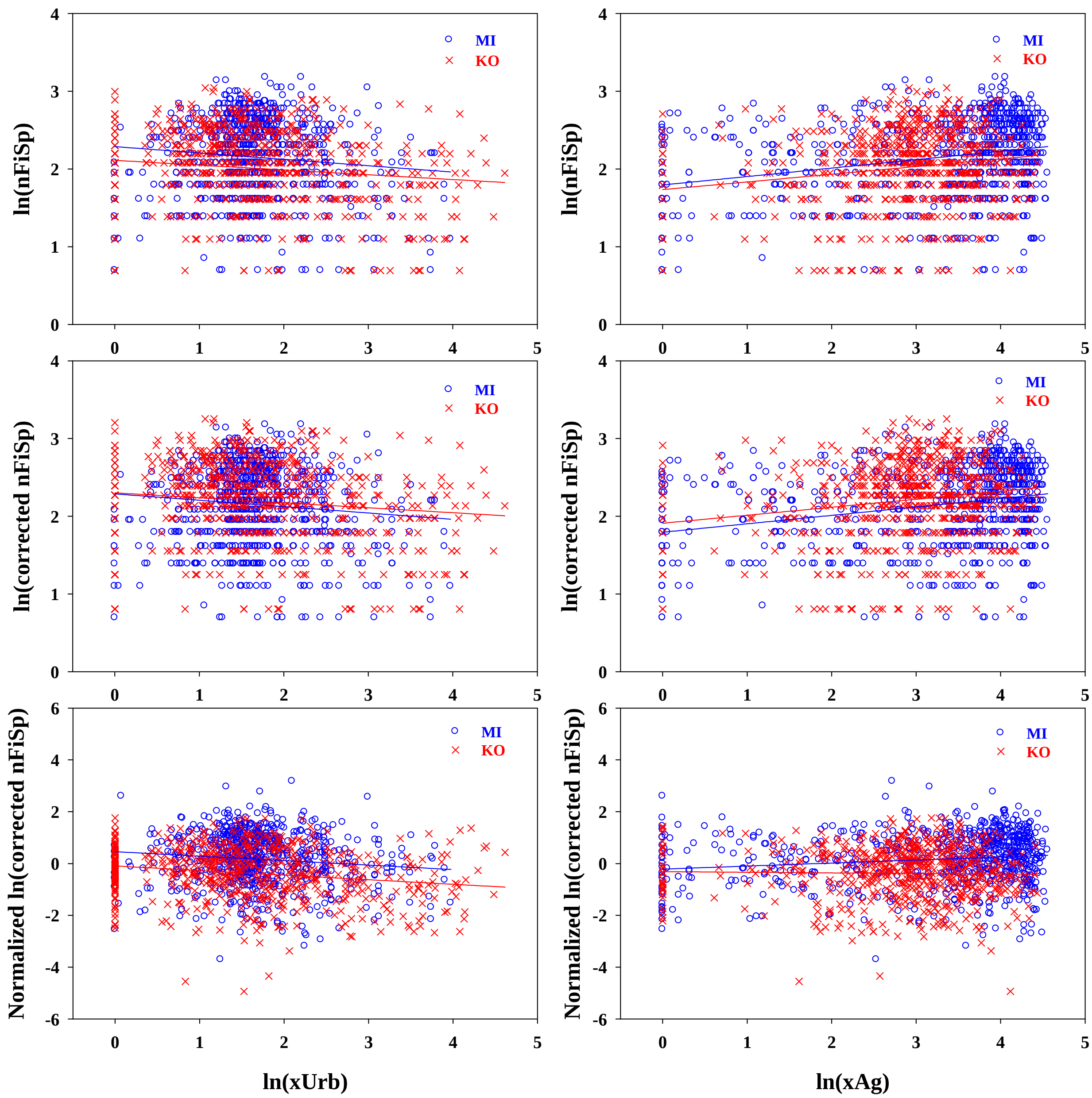


Figure 6.9. Scatter plots of Number of intolerant fish species (nFiInt) against urban and agricultural LUs Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data).

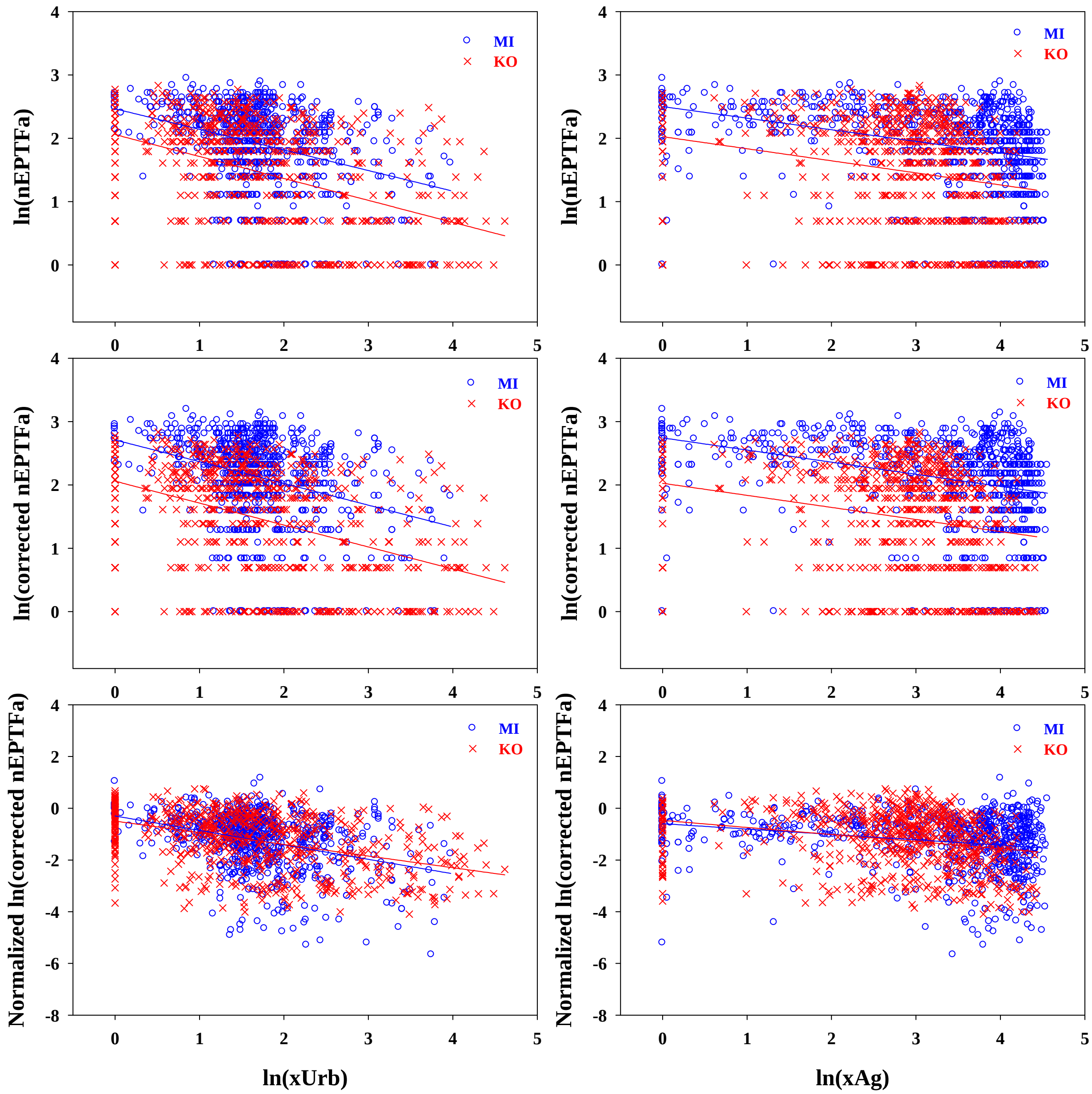


Figure 6.10. Scatter plots of Number of EPT families (nEPTFa) against urban and agricultural LUs Michigan and Korea (large river site excluded) for three types of datasets (raw, method corrected, and normalized method-corrected data).

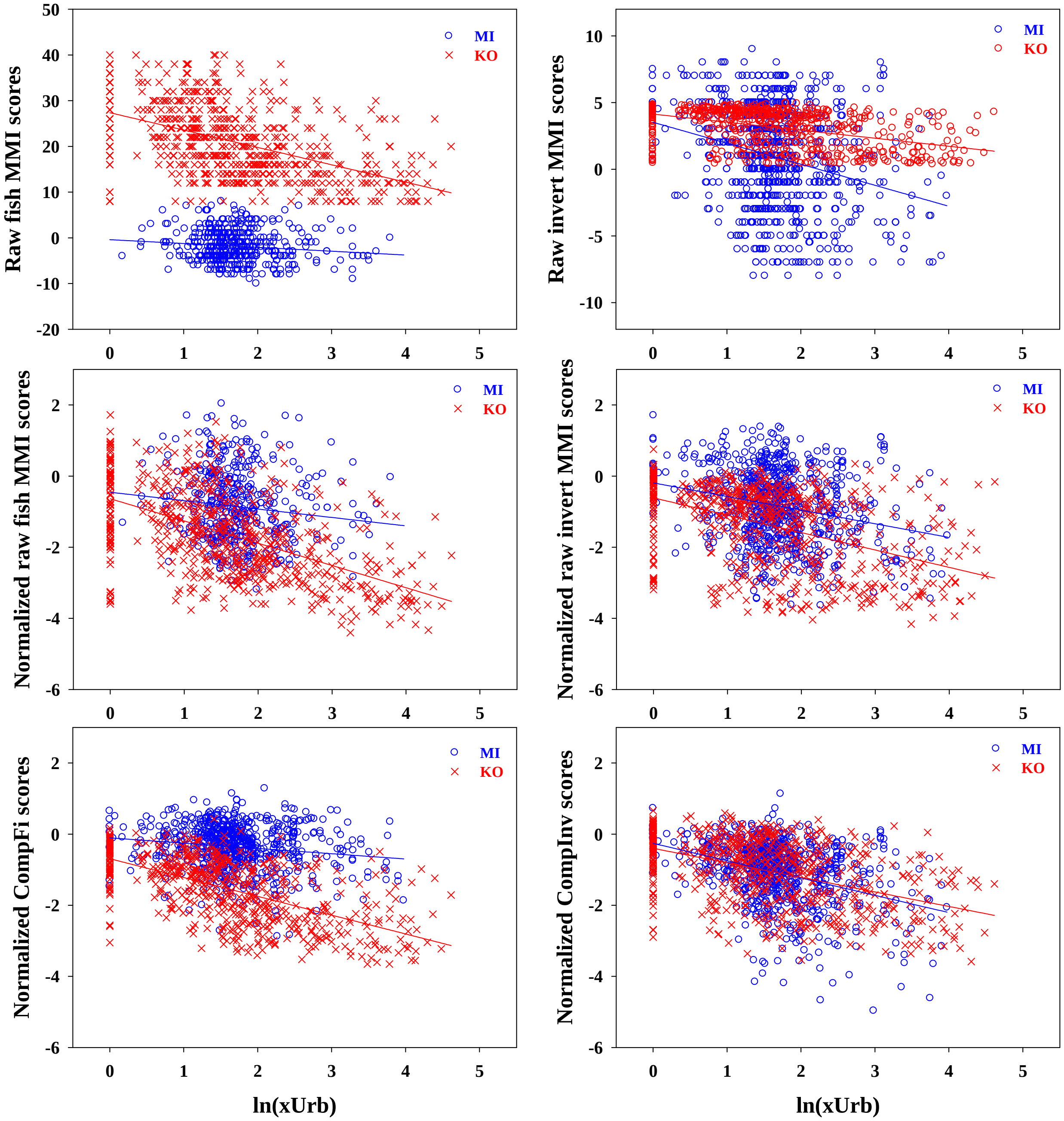


Figure 6.11. Scatter plots of stream health assessment scores against urban LU gradients. Assessment scores were produced by regional raw MMIs (top row), normalized regional raw MMIs (middle row), and normalized composite MMIs (bottom row) for fish and benthic macroinvertebrates. Blue and circles indicate Michigan samples and red and x mark symbols indicate Korean samples (large river sites excluded).

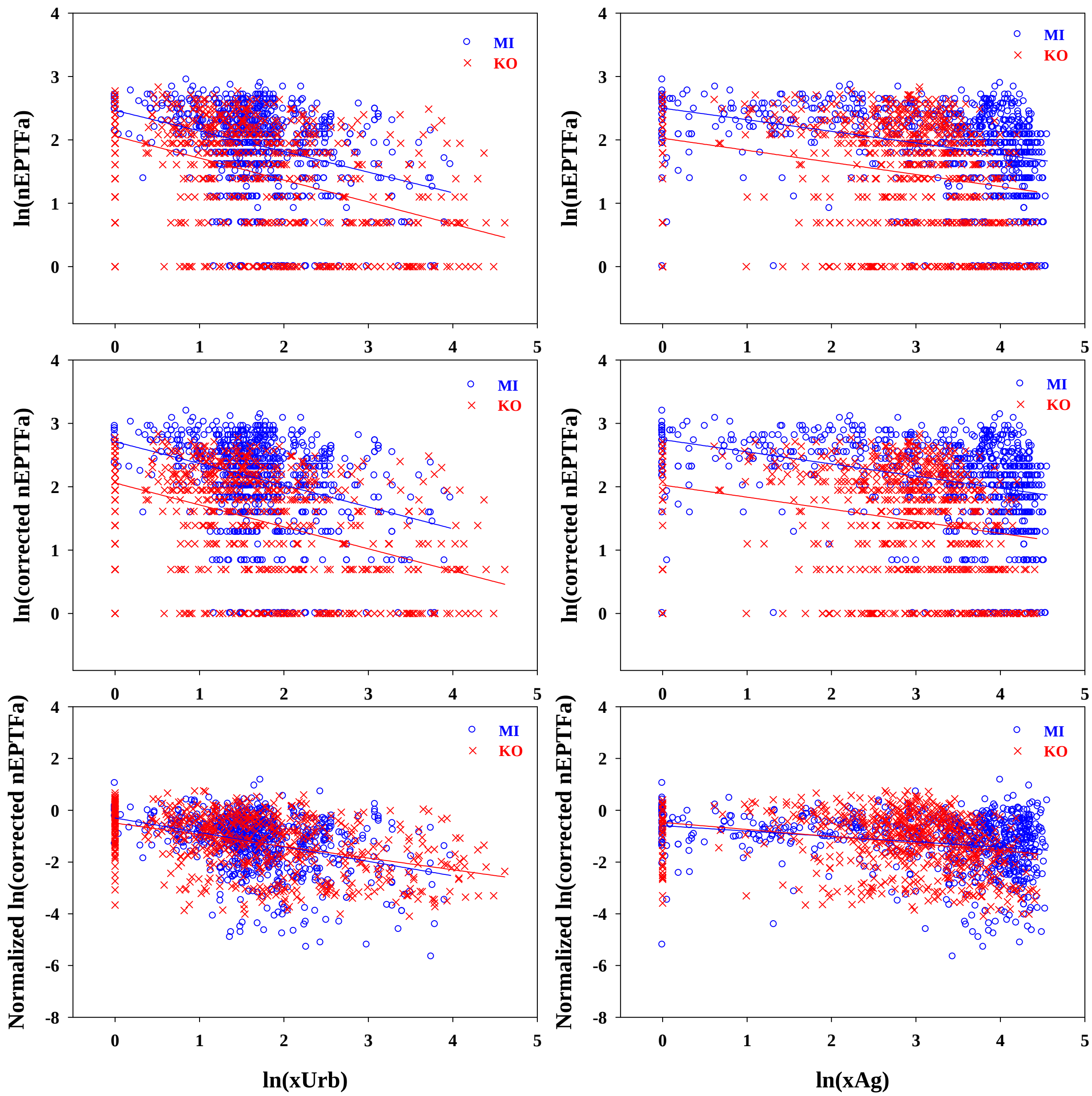


Figure 6.12. Scatter plots of stream health assessment scores against agricultural LU gradients. Assessment scores were produced by regional raw MMIs (top row), normalized regional raw MMIs (middle row), and normalized composite MMIs (bottom row) for fish and benthic macroinvertebrates. Blue and circles indicate Michigan samples and red and x mark symbols indicate Korean samples (large river sites excluded).

Appendix 6.1. Summary statistics of landscape variables for Korean data (large river sites were included). SD indicates standard deviation.

Variables	n	Mean	Median	SD	Min	Max
Korean region (large river sites included)						
Drainage Area; km^2	684	1110.5	141.3	3285.1	0.6	25637.4
Average wetted width; m	684	63.7	30.0	113.8	0.2	1500.0
Average depth; m	684	0.64	0.30	1.14	0.00	15.00
Water temperature; ºC	684	15.5	15.9	4.5	1.0	27.6
Percent of urban land use (xUrb)	684	8.71	3.67	13.80	0.00	100.0
Percent of agricultural land use (xAg)	684	22.08	18.76	16.30	0.00	83.10
High-flow yield (Q10Y)	684	0.079	0.070	0.050	0.011	0.447
Median-flow yield (Q50Y)	684	0.027	0.021	0.022	0.004	0.297
Low-flow yield (Q90Y)	684	0.014	0.010	0.013	0.002	0.094

Appendix 6.2. Summary statistics of raw data of biological indicator metrics (fish and benthic macroinvertebrates) used for the study. SD indicates standard deviation.

Variables (raw data)	n	Mean	Median	SD	Min	Max
Michigan data						
Number of total fish species (nFiSp)	746	8.6	8.0	4.4	1.0	23.0
Number of intolerant fish species (nFiInt)	746	2.4	2.0	1.7	0.0	9.0
Percent of tolerant fish individuals (pFiTol)	746	37.9	34.2	29.4	0.0	100.0
Percent of omnivorous fish individuals (pFiOmn)	746	28.9	22.9	25.9	0.0	100.0
Percent of insectivorous fish individuals (pFiIns)	746	41.6	41.8	24.4	0.0	100.0
Number of total invertebrate families (nFaInv)	784	21.3	20.6	7.0	3.0	43.0
Number of EPT families (nEPTFa)	784	6.9	7.0	3.8	0.0	18.0
Percent of total individuals that were EPT (pEPTIn)	784	35.0	35.3	19.3	0.0	89.6
MBI Biotic Index (MBI) score	784	5.5	5.5	0.8	3.6	7.9
Korean data with large river sites included						
Number of total fish species (nFiSp)	663	7.5	7.0	3.9	1.0	20.0
Number of intolerant fish species (nFiInt)	663	2.0	1.0	2.2	0.0	10.0
Percent of tolerant fish individuals (pFiTol)	663	48.5	46.8	35.7	0.0	100.0
Percent of omnivorous fish individuals (pFiOmn)	663	48.2	49.3	32.9	0.0	100.0
Percent of insectivorous fish individuals (pFiIns)	663	43.0	39.6	32.9	0.0	100.0
Number of total invertebrate families (nFaInv)	677	10.7	10.0	5.6	1.0	27.0
Number of EPT families (nEPTFa)	677	5.0	5.0	4.0	0.0	16.0
Percent of total individuals that were EPT (pEPTIn)	677	37.0	31.6	32.6	0.0	100.0
MBI Biotic Index (MBI) score	673	5.6	5.5	1.6	0.7	10.0
Korean data with large river sites excluded						
Number of total fish species (nFiSp)	612	7.4	7.0	3.9	1.0	20.0
Number of intolerant fish species (nFiInt)	612	2.1	1.0	2.3	0.0	10.0
Percent of tolerant fish individuals (pFiTol)	612	47.8	45.8	36.0	0.0	100.0
Percent of omnivorous fish individuals (pFiOmn)	612	48.4	49.7	33.5	0.0	100.0
Percent of insectivorous fish individuals (pFiIns)	612	43.5	39.7	33.5	0.0	100.0
Number of total invertebrate families (nFaInv)	627	10.9	10.0	5.6	1.0	27.0
Number of EPT families (nEPTFa)	627	5.1	5.0	4.0	0.0	16.0
Percent of total individuals that were EPT (pEPTIn)	627	37.4	33.9	32.6	0.0	100.0
MBI Biotic Index (MBI) score	624	5.6	5.5	1.6	0.7	10.0

Appendix 6.3. Summary statistics of method corrected data of biological indicator metrics (fish and benthic macroinvertebrates) used for the study. All raw data of each metric were corrected by statistical equations (Table 6.2). SD indicates standard deviation.

Variables (method corrected)	n	Mean	Median	SD	Min	Max
Michigan data						
Number of total fish species (nFiSp)	746	8.6	8.0	4.4	1.0	23.0
Number of intolerant fish species (nFiInt)	746	2.4	2.0	1.7	0.0	9.0
Percent of tolerant fish individuals (pFiTol)	746	37.9	34.2	29.4	0.0	100.0
Percent of omnivorous fish individuals (pFiOmn)	746	28.9	22.9	25.9	0.0	100.0
Percent of insectivorous fish individuals (pFiIns)	746	41.6	41.8	24.4	0.0	100.0
Number of total invertebrate families (nFaInv)	784	27.7	26.8	9.2	3.9	55.9
Number of EPT families (nEPTFa)	784	8.9	9.1	5.0	0.0	23.3
Percent of total individuals that were EPT (pEPTIn)	784	34.2	34.5	18.8	0.0	87.4
MBI Biotic Index (MBI) score	784	5.2	5.2	0.7	3.4	7.4
Korean data with large river sites included						
Number of total fish species (nFiSp)	663	9.3	8.7	4.9	1.2	24.9
Number of intolerant fish species (nFiInt)	663	2.6	1.3	2.8	0.0	12.7
Percent of tolerant fish individuals (pFiTol)	663	43.5	41.9	32.0	0.0	89.7
Percent of omnivorous fish individuals (pFiOmn)	663	47.1	48.2	32.2	0.0	97.8
Percent of insectivorous fish individuals (pFiIns)	663	38.2	35.2	29.2	0.0	88.7
Number of total invertebrate families (nFaInv)	677	10.7	10.0	5.6	1.0	27.0
Number of EPT families (nEPTFa)	677	5.0	5.0	4.0	0.0	16.0
Percent of total individuals that were EPT (pEPTIn)	677	37.0	31.6	32.6	0.0	100.0
MBI Biotic Index (MBI) score	673	5.6	5.5	1.6	0.7	10.0
Korean data with large river sites excluded						
Number of total fish species (nFiSp)	612	9.2	8.7	4.9	1.2	24.9
Number of intolerant fish species (nFiInt)	612	2.7	1.3	2.9	0.0	12.7
Percent of tolerant fish individuals (pFiTol)	612	42.8	41.1	32.3	0.0	89.7
Percent of omnivorous fish individuals (pFiOmn)	612	47.3	48.6	32.7	0.0	97.8
Percent of insectivorous fish individuals (pFiIns)	612	38.6	35.2	29.7	0.0	88.7
Number of total invertebrate families (nFaInv)	627	10.9	10.0	5.6	1.0	27.0
Number of EPT families (nEPTFa)	627	5.1	5.0	4.0	0.0	16.0
Percent of total individuals that were EPT (pEPTIn)	627	37.4	33.9	32.6	0.0	100.0
MBI Biotic Index (MBI) score	624	5.6	5.5	1.6	0.7	10.0

Appendix 6.4. Summary statistics of Independent samples t-tests of biological indicator metrics between Michigan and Korean data. CI and df indicate confidence interval and degree of freedom, respectively.

Metrics	Mean difference	Std. error difference	95% CI of the difference		t	df	Sig. (2-tailed)
			Lower	Upper			
A. Raw data (Michigan versus Korea with large river sites excluded)							
Number of total fish species (nFiSp)	1.203	0.228	0.756	1.650	5.280	1356	.000
Number of intolerant fish species (nFiInt)	0.304	0.108	0.091	0.516	2.805	1356	.005
Percent of tolerant fish individuals (pFiTol)	-9.856	1.775	-13.339	-6.374	-5.553	1356	.000
Percent of omnivorous fish individuals (pFiOmn)	-19.494	1.611	-22.653	-16.335	-12.104	1356	.000
Percent of insectivorous fish individuals (pFiIns)	-1.916	1.573	-5.003	1.170	-1.218	1356	.223
Number of total invertebrate families (nFaInv)	10.445	0.346	9.767	11.123	30.219	1409	.000
Number of EPT families (nEPTFa)	1.818	0.211	1.405	2.231	8.632	1409	.000
Percent of total individuals that were EPT (pEPTIn)	-2.429	1.395	-5.165	0.307	-1.741	1409	.082
MBI Biotic Index (MBI) score	-0.103	0.065	-0.230	0.024	-1.584	1406	.113
B. Method corrected data (Michigan versus Korea with large river sites excluded)							
Number of total fish species (nFiSp)	-0.650	0.254	-1.147	-0.153	-2.563	1332	.010
Number of intolerant fish species (nFiInt)	-0.273	0.127	-0.523	-0.024	-2.151	1335	.032
Percent of tolerant fish individuals (pFiTol)	-5.147	1.685	-8.451	-1.842	-3.055	1346	.002
Percent of omnivorous fish individuals (pFiOmn)	-18.587	1.597	-21.720	-15.454	-11.638	1346	.000
Percent of insectivorous fish individuals (pFiIns)	3.162	1.474	.271	6.053	2.146	1346	.032
Number of total invertebrate families (nFaInv)	16.163	0.455	15.272	17.055	35.558	1430	.000
Number of EPT families (nEPTFa)	3.752	0.245	3.271	4.234	15.295	1427	.000
Percent of total individuals that were EPT (pEPTIn)	-3.268	1.385	-5.985	-0.552	-2.360	1409	.018
MBI Biotic Index (MBI) score	-0.453	0.064	-0.578	-0.328	-7.094	1406	.000

Appendix 6.5. Coefficients and their statistical significance from GLM ANCOVA tests of biological indicator metrics with landscape variables between Michigan and Korean (large river site excluded) datasets. Region (Michigan and Korea) was used as fixed factor and landscape variables (drainage area, xUrb, and xAg) were used as covariates. Bold indicates significance at $p \le 0.05$ and bold and italics indicate significance at $p \le 0.01$. Coefficients for Region, Region*ln(xUrb), and Region*ln(xAg) were summarized for Korean data and coefficients of these for Michigan datacould be calculated by multiplying by -1 to coefficients of Korean data.

Covariates	Biological indicator metrics								
	ln(nFiSp)	ln(nFiInt)	ln(pFiTol)	ln(pFiOmn)	ln(pFiIns)	ln(nFaInv)	ln(nEPTFa)	ln(pEPTIn)	ln(MBI)
A. *Raw data (Michigan versus Korea (large river sites excluded))*									
Region	0.01198	***0.2506***	***-0.4355***	***-0.61***	***1.208***	***-0.2934***	**-0.1532**	-0.0009773	0.002386
ln(Drainage area)	***0.1839***	***0.1915***	**0.1809**	***0.3213***	**0.1486**	***0.09901***	***0.1383***	***0.1877***	-0.008811
Region*ln(Drainage area)	**-0.0194**	**-0.02152**	-0.009457	***0.07685***	***-0.09507***	**-0.01496**	-0.003276	0.0424	***-0.01055***
ln(xUrb)	-0.003695	0.03693	0.2201	***0.3934***	***-0.5155***	***-0.1507***	***-0.252***	***-0.3086***	***0.07499***
Region*ln(xUrb)	0.01699	***-0.1189***	***0.3461***	***0.2705***	***-0.2659***	**-0.03059**	**-0.06018**	***-0.1271***	***0.02804***
ln(Drainage area)*ln(xUrb)	-0.005503	***-0.05501***	0.0326	-0.02573	***0.07057***	-0.005158	-0.001301	-0.01335	-0.003069
ln(xAg)	***0.1659***	***-0.159***	***0.8064***	***0.826***	0.05303	-0.006824	***-0.1515***	***-0.3149***	***0.05027***
Region*Ln(xAg)	-0.02092	***-0.04932***	0.02648	0.06375	***-0.1929***	0.008277	-0.0085	***-0.1209***	0.003854
ln(Drainage area)*ln(xAg)	-0.007223	0.007153	**-0.0448**	***-0.07867***	-0.02989	-0.009422	-0.004559	0.009391	-0.0004472
ln(xUrb)*ln(xAg)	***-0.03348***	-0.01755	**-0.06131**	-0.06231	**-0.05947**	0.005233	-0.005607	-0.0147	-0.003178
B. *Method corrected data (Michigan versus Korea (large river sites excluded))*									
Region	0.09272	***0.3517***	***-0.4485***	***-0.6136***	***1.149***	***-0.4193***	***-0.269***	0.01054	0.02903
ln(Drainage area)	***0.1878***	***0.2031***	**0.1755**	***0.3198***	**0.1487**	***0.09963***	***0.1377***	***0.1879***	-0.008889
Region*ln(Drainage area)	-0.01701	-0.01794	-0.01117	***0.07641***	***-0.09704***	**-0.01539**	-0.006813	0.04266	***-0.01056***
ln(xUrb)	-0.004212	0.04055	0.2096	***0.3911***	***-0.5043***	***-0.1515***	***-0.2568***	***-0.3082***	***0.07476***
Region*ln(xUrb)	0.01539	***-0.1381***	***0.3417***	***0.2698***	***-0.2583***	**-0.02968**	**-0.05365**	***-0.1276***	***0.02818***
ln(Drainage area)*ln(xUrb)	-0.005452	***-0.05984***	0.03303	-0.02546	***0.06891***	-0.00516	-0.001189	-0.01336	-0.003048
ln(xAg)	***0.1698***	***-0.1653***	***0.794***	***0.824***	0.05308	-0.006708	***-0.1592***	***-0.3142***	***0.04998***
Region*Ln(xAg)	-0.01965	***-0.06034***	0.02048	0.06282	***-0.1908***	0.008501	-0.004616	***-0.1211***	0.004061
ln(Drainage area)*ln(xAg)	-0.007716	0.007261	**-0.04388**	***-0.0785***	-0.02964	-0.009486	-0.003162	0.00925	-0.0004262
ln(xUrb)*ln(xAg)	***-0.03396***	-0.02042	-0.0594	-0.06204	**-0.05863**	0.005188	-0.006621	-0.01464	-0.003184

Appendix 6.6. Coefficients and their statistical significance from GLM ANCOVA tests of biological indicator metrics with landscape variables between Michigan and Korean (large river site included) datasets. Region (Michigan and Korea) was used as fixed factor and landscape variables (drainage area, xUrb, and xAg) were used as covariates. Bold indicates significance at $p \leq 0.05$ and bold and italics indicate significance at $p \leq 0.01$. Coefficients for Region, Region*ln(xUrb), and Region*ln(xAg) were summarized for Korean data and coefficients of these for Michigan data could be calculated by multiplying by -1 to coefficients of Korean data.

Covariates	Biological indicator metrics								
	ln(nFiSp)	ln(nFiInt)	ln(pFiTol)	ln(pFiOmn)	ln(pFiIns)	ln(nFaInv)	ln(nEPTFa)	ln(pEPTIn)	ln(MBI)
A. Raw data (Michigan versus Korea (large river sites included))									
Region	0.0843	***0.3353***	***-0.4002***	***-0.5108***	***1.217***	***-0.2459***	-0.05881	0.1767	-0.01523
ln(Drainage area)	***0.1637***	***0.1351***	***0.2512***	***0.3656***	**0.1252**	***0.08888***	***0.1289***	***0.1872***	-0.005252
Region*ln(Drainage area)	***-0.04116***	***-0.04962***	-0.01523	**0.05149**	***-0.09995***	***-0.03105***	***-0.03402***	-0.01181	-0.005145
ln(xUrb)	-0.0199	-0.01547	***0.329***	***0.4643***	***-0.5277***	***-0.1259***	***-0.1933***	**-0.2138**	***0.07471***
Region*ln(xUrb)	0.01768	***-0.1187***	***0.3514***	***0.2757***	***-0.2655***	-0.02729	**-0.05401**	***-0.1175***	***0.0274***
ln(Drainage area)*ln(xUrb)	0.000606	***-0.03706***	-0.002111	-0.04761	***0.07468***	-0.0128	-0.01904	-0.04238	-0.002676
ln(xAg)	***0.1878***	***-0.1412***	***0.8153***	***0.8613***	0.04456	-0.01039	***-0.1534***	***-0.3001***	***0.04851***
Region*Ln(xAg)	-0.01406	**-0.0372**	0.02105	0.06572	***-0.1889***	0.01424	0.001916	***-0.1058***	0.002263
ln(Drainage area)*ln(xAg)	-001104	0.006299	**-0.04997**	***-0.0888***	-0.02626	-0.00711	-0.001535	0.008398	-0.00007953
ln(xUrb)*ln(xAg)	***-0.03455***	-0.02083	-0.05487	-0.05744	**-0.06053**	0.006636	***-0.003109***	-0.009138	-0.003981
B. Method corrected data (Michigan versus Korea (large river sites included))									
Region	***0.1676***	***0.4429***	***-0.4149***	***-0.515***	***1.157***	***-0.3718***	**-0.1742**	0.1882	0.01142
ln(Drainage area)	***0.1669***	***0.141***	***0.2448***	***0.364***	**0.1258**	***0.08947***	***0.1286***	***0.1873***	-0.005318
Region*ln(Drainage area)	***-0.03955***	***-0.04831***	-0.01653	**0.05119**	***-0.1016***	***-0.03148***	***-0.03759***	-0.01154	-0.005161
ln(xUrb)	-0.02052	-0.01722	***0.3171***	***0.4619***	***-0.5164***	***-0.1267***	***-0.198***	**-0.2135**	***0.0745***
Region*ln(xUrb)	0.01615	***-0.1379***	***0.3469***	***0.275***	***-0.2579***	-0.02637	-0.04751	***-0.1179***	***0.02755***
ln(Drainage area)*ln(xUrb)	0.0007213	***-0.04006***	-0.001243	-0.04732	***0.07297***	-0.0128	-0.01895	-0.04239	-0.002661
ln(xAg)	***0.1921***	***-0.1465***	***0.8022***	***0.859***	0.04468	-0.01029	***-0.1607***	***-0.2995***	***0.04823***
Region*Ln(xAg)	-0.01254	***-0.04708***	0.01507	0.06479	***-0.187***	0.01446	0.005736	***-0.106***	0.002469
ln(Drainage area)*ln(xAg)	-0.01161	0.006518	**-0.04884**	***-0.08856***	-0.02606	-0.007168	-0.000266	0.008271	-0.00005982
ln(xUrb)*ln(xAg)	***-0.03503***	-0.02404	-0.05307	-0.05718	**-0.05968**	0.006591	-0.004108	-0.009084	-0.003986

Appendix 6.7. Coefficients and their statistical significance from GLM ANCOVA tests of normalized scores of biological indicator metrics with urban and agricultural LUs between Michigan and Korean (large river site included) datasets. Region (Michigan and Korea) was used as fixed factor and LU stressors (xUrb and xAg) were used as covariates. Bold indicates significance at $p \leq 0.05$ and bold and italics indicate significance at $p \leq 0.01$. Coefficients for Region, Region*ln(xUrb), and Region*ln(xAg) were summarized for Korean data and coefficients of these for Michigan data could be calculated by multiplying by -1 to coefficients of Korean data.

Covariates	Biological indicator metrics								
	ln(nFiSp)	ln(nFiInt)	ln(pFiTol)	ln(pFiOmn)	ln(pFiIns)	ln(nFaInv)	ln(nEPTFa)	ln(pEPTIn)	ln(MBI)
A. *Normalized raw data (Michigan versus Korea (large river sites included))*									
Region	0.07697	0.06028	0.01508	0.04461	-0.02589	-0.0201	0.002418	0.002399	0.001585
ln(xUrb)	-0.07308	***-0.2789***	***0.2593***	***0.2422***	***-0.2788***	***-0.3872***	***-0.3816***	***-0.3015***	***0.3976***
Region*ln(xUrb)	-0.03168	***0.2635***	***-0.3089***	***-0.1756***	***0.1861***	-0.02794	-0.02013	0.01824	-0.006821
ln(xAg)	***0.2457***	-0.0819	***0.4117***	***0.3228***	-0.0528	-0.06863	**-0.1094**	***-0.1216***	***0.1842***
Region*ln(xAg)	0.02212	***0.1473***	-0.0292	***-0.08543***	***0.1493***	-0.03539	0.03011	***0.09641***	-0.002481
ln(xUrb)*ln(xAg)	***-0.06989***	***-0.07014***	-0.04087	**-0.05159**	-0.03962	-0.008907	-0.0389	-0.03373	-0.004567
B. *Normalized method corrected data (Michigan versus Korea (large river sites included))*									
Region	0.07683	0.06119	0.01475	0.04134	-0.02584	-0.02012	0.002608	0.0024	0.001603
ln(xUrb)	-0.07388	***-0.2769***	***0.2591***	***0.2421***	***-0.2791***	***-0.3875***	***-0.3783***	***-0.3017***	***0.3975***
Region*ln(xUrb)	-0.03166	***0.2648***	***-0.3107***	***-0.1757***	***0.1866***	-0.02803	-0.01803	0.01808	-0.006785
ln(xAg)	***0.2453***	-0.07792	***0.4106***	***0.3225***	-0.05423	-0.06842	**-0.1046**	***-0.1218***	***0.184***
Region*ln(xAg)	0.02193	***0.1456***	-0.02941	***-0.08561***	***0.1509***	-0.03505	0.03399	***0.09623***	-0.002541
ln(xUrb)*ln(xAg)	***-0.06955***	***-0.07159***	-0.03998	**-0.05145**	-0.03976	-0.008836	-0.03943	-0.03373	-0.004513

Appendix 6.8. Multiple linear regression models of biological indicator metrics for Michigan data set. Bold indicates significance at p ≤ 0.05 and bold and italics indicate significance at p ≤ 0.01.

Covariates	Biological indicator metrics								
	ln(nFiSp)	ln(nFiInt)	ln(pFiTol)	ln(pFiOmn)	ln(pFiIns)	ln(nFaInv)	ln(nEPTFa)	ln(pEPTIn)	ln(MBI)
A. *Michigan region: raw data*									
R^2 (%)	35.4	29.0	27.0	19.6	16.7	22.2	36.0	27.4	38.7
Constant	***0.51173***	***3.04466***	***-2.4977***	***-3.94568***	***1.94777***	***2.53264***	***4.13222***	***7.46767***	***1.2496***
ln(Drainage area)	***0.146074***	*0.144511*	0.0653956	**-0.0766196**	***0.25418***	***0.0697973***	***0.121176***	***0.184576***	***-0.00650826***
ln(xUrb)	***-0.114192***	***-0.0970936***	**-0.17656**		***-0.152761***	***-0.13125***	***-0.242752***	***-0.27673***	***0.0353361***
ln(xAg)	***0.0680155***		***0.393127***	***0.19205***		***-0.0355811***	***-0.070032***	**-0.0606591**	***0.0161977***
ln(water temperature)	**0.133092**	***-0.317016***		***0.99961***			***-0.229785***	***-0.402126***	***0.0620031***
ln(site slope)		***0.0952999***	***-0.372855***	**-0.169997**	***-0.173985***	**0.0291499**	***0.0855778***	***0.273371***	***-0.0244332***
ln(Q10Y)			***-0.427049***			***-0.19905***		**0.288407**	
ln(Q50Y)				***-0.444247***					
ln(Q90Y)	***-0.0918507***	***0.127008***			**0.0828547**		***0.128197***		***-0.0284357***
ln(number of dams)	***0.0779854***								
B. *Michigan region: method corrected data*									
R^2 (%)	35.4	29.0	27.0	19.6	16.7	22.2	35.5	27.5	38.7
Constant	***0.51173***	***3.04466***	***-2.4977***	***-3.94568***	***1.94777***	***2.7767***	***4.49288***	***7.43006***	***1.20111***
ln(Drainage area)	***0.146074***	***0.144511***	0.0653956	**-0.0766196**	***0.25418***	***0.0705621***	***0.12837***	***0.183896***	***-0.00645228***
ln(xUrb)	***-0.114192***	***-0.0970936***	**-0.17656**		***-0.152761***	***-0.133132***	***-0.257765***	***-0.275792***	***0.0349737***
ln(xAg)	***0.0680155***		***0.393127***	***0.19205***		***-0.0358762***	***-0.0708254***	**-0.0606796**	***0.0160165***
ln(water temperature)	**0.133092**	***-0.317016***		***0.99961***			***-0.244676***	***-0.40065***	***0.0613356***
ln(site slope)		***0.0952999***	***-0.372855***	**-0.169997**	***-0.173985***	**0.0296308**	***0.0928958***	***0.27235***	***-0.0241759***
ln(Q10Y)			***-0.427049***			***-0.201914***		**0.28714**	
ln(Q50Y)				***-0.444247***					
ln(Q90Y)	***-0.0918507***	***0.127008***			**0.0828547**		***0.1367***		***-0.028149***
ln(number of dams)	***0.0779854***								

Appendix 6.9. Multiple linear regression models of biological indicator metrics for Korean data set with large river sites excluded. Bold indicates significance at p ≤ 0.05 and bold and italics indicate significance at p ≤ 0.01.

Covariates	Biological indicator metrics								
	ln(nFiSp)	ln(nFiInt)	ln(pFiTol)	ln(pFiOmn)	ln(pFiIns)	ln(nFaInv)	ln(nEPTFa)	ln(pEPTIn)	ln(MBI)
A. Korean region: raw data (large river sites excluded)									
R^2 (%)	26.3	45.9	39.7	26.4	34.2	17.1	30.4	24.9	20.2
Constant	0.401871	***2.05334***	***-1.04681***	***0.640272***	***5.53628***	***2.62235***	***3.26994***	***5.39404***	***1.6583***
ln(Drainage area)	***0.128993***	***0.236659***	**-0.0950364**	***0.165888***	***0.247567***	***0.0587684***	***0.230245***	***0.407721***	***-0.0338956***
ln(xUrb)	***-0.079593***	***-0.393951***	***0.589382***	***0.40284***	***-0.665438***	***-0.181628***	***-0.312007***	***-0.544448***	***0.0865388***
ln(xAg)		***-0.176502***	***0.453476***	***0.455706***	***-0.306202***		***-0.128074***	***-0.358115***	***0.0434678***
ln(water temperature)	**0.129061**						***-0.283464***	**-0.390788**	
ln(site slope)	***-0.0450281***		**-0.111839**			**0.0454843**	***0.0800071***		
ln(Q10Y)		***0.434991***	***-0.708398***		***0.578985***		***0.274842***	***0.563564***	-0.0391088
ln(Q50Y)	***-0.143987***								
ln(number of dams)	***-0.363097***	***-0.3391***		-0.33831	***-0.487281***		***-0.323795***	**-0.528467**	
B. Korean region: method corrected data (large river sites excluded)									
R^2 (%)	26.3	45.9	39.3	26.4	34.4	17.1	30.4	24.9	20.2
Constant	**0.529148**	***2.30892***	***-0.77162***	***0.633924***	***5.4134***	***2.62235***	***3.26994***	***5.39404***	***1.6583***
ln(Drainage area)	***0.133529***	***0.257232***	-0.0749773	***0.164634***	***0.240871***	***0.0587684***	***0.230245***	***0.407721***	***-0.0338956***
ln(xUrb)	***-0.0823575***	***-0.440219***	***0.611409***	***0.401648***	***-0.651046***	***-0.181628***	***-0.312007***	***-0.544448***	***0.0865388***
ln(xAg)		***-0.195186***	***0.472554***	***0.453941***	***-0.30256***		***-0.128074***	***-0.358115***	***0.0434678***
ln(water temperature)	**0.133365**						***-0.283464***	**-0.390788**	
ln(site slope)	***-0.0469718***					**0.0454843**	***0.0800071***		
ln(Q10Y)		***0.476824***	***-0.743432***		***0.572171***		***0.274842***	***0.563564***	-0.0391088
ln(Q50Y)	***-0.14936***								
ln(number of dams)	***-0.376008***	***-0.371055***		-0.335665	***-0.471403***		***-0.323795***	**-0.528467**	

Appendix 6.10. Multiple linear regression models of biological indicator metrics for Korean data set with large river sites included. Bold indicates significance at p ≤ 0.05 and bold and italics indicate significance at p ≤ 0.01.

Covariates	Biological indicator metrics								
	ln(nFiSp)	ln(nFiInt)	ln(pFiTol)	ln(pFiOmn)	ln(pFiIns)	ln(nFaInv)	ln(nEPTFa)	ln(pEPTIn)	ln(MBI)
A. Korean region: raw data (large river sites included)									
R^2 (%)	16.9	43.7	37.9	24.5	31.8	17.6	30.3	25.2	18.6
Constant	***1.29144***	***2.0328***	***-0.863519***	***0.691263***	***5.6435***	***2.57662***	***3.19357***	***5.32373***	***1.60303***
ln(Drainage area)	***0.0782991***	***0.218677***	**-0.0966052**	***0.17564***	***0.208761***	***0.0673006***	***0.230156***	***0.415981***	***-0.0284931***
ln(xUrb)	***-0.0936728***	***-0.379859***	***0.56439***	***0.38475***	***-0.63635***	***-0.185938***	***-0.32185***	***-0.559032***	***0.0862003***
ln(xAg)	**0.0680155**	***-0.161805***	***0.437866***	***0.436982***	***-0.311881***		***-0.115507***	***-0.339449***	***0.0390179***
ln(water temperature)							***-0.253015***	-0.339711	
ln(site slope)	***-0.0596839***		***-0.106667***			**0.0400202**	***0.0730213***		
ln(Q10Y)		***0.420167***	***-0.68286***		***0.580432***		***0.296977***	***0.609872***	**-0.0570871**
ln(number of dams)		***-0.391187***		***-0.485168***		***-0.226492***	***-0.461334***	***-0.763313***	
B. Korean region: method corrected data (large river sites included)									
R^2 (%)	17.0	43.7	38.0	24.5	32.0	17.6	30.3	25.2	18.6
Constant	***1.45057***	***2.28623***	***-0.865212***	***0.684685***	***5.51339***	***2.57662***	***3.19357***	***5.32373***	***1.60303***
ln(Drainage area)	***0.0811684***	***0.237915***	**-0.0957932**	***0.174357***	***0.203629***	***0.0673006***	***0.230156***	***0.415981***	***-0.0284931***
ln(xUrb)	***-0.0970271***	***-0.424668***	***0.556116***	***0.383597***	***-0.622642***	***-0.185938***	***-0.32185***	***-0.559032***	***0.0862003***
ln(xAg)	**0.048892**	***-0.179021***	***0.428998***	***0.435312***	***-0.308255***		***-0.115507***	***-0.339449***	***0.0390179***
ln(water temperature)							***-0.253015***	-0.339711	
ln(site slope)	***-0.0619447***		**-0.103064**			**0.0400202**	***0.0730213***		
ln(Q10Y)		***0.461328***	***-0.670209***		***0.572547***		***0.296977***	***0.609872***	**-0.0570871**
ln(number of dams)		***-0.425371***		***-0.482145***		***-0.226492***	***-0.461334***	***-0.763313***	

Appendix 6.11. Multiple linear regression models of regional multimetrics. Bold indicates significance at p ≤ 0.05 and bold and italics indicate significance at p ≤ 0.01.

Variables	Korean regional MMIs		Michigan regional MMIs	
	KIBI	*KSI*	*P51Fi*	*P51Inv*
R^2 (%)	38.8	29.5	6.3	25.6
Constant	***39.3369***	***0.653916***	***-0.895128***	***10.2632***
ln(Drainage area)	***1.41893***	***-0.28186***	***0.443138***	***0.713535***
ln(xUrb)	***-3.73595***	***0.546502***	**-0.727253**	***-1.09932***
ln(xAg)	***-1.40919***	***0.266573***	**-0.456628**	**-0.278505**
ln(water temperature)				***-1.76534***
ln(site slope)	***0.685186***	***-0.146264***		***1.30707***
ln(Q10Y)	***4.02555***			**-1.23617**
ln(number of dams)	***-3.17926***	***0.499477***		

Literature cited

Allan, J. D. and L. B. Johnson. 1997. Cathcment scale analysis of aquatic ecosystems. Freshwater Biology. 37: 107-111.

Allan, J. D., P. B. McIntyre, S. D. P. Smith, B. S. Halpern, G. L. Boyer, A. Buchsbaum, G. A. Burton Jr., L. M. Campbell, W. L. Chadderton, J. J. H. Ciborowski, P. J. Doran, T. Eder, D. M. Infante, L. B. Johnson, C. A. Joseph, A. L. Marino, A. Prusevich, J. Read, J. B. Rose, E. S. Rutherford, S. P. Sowa, A. D. Steinman. 2013. Joint analysis of stressors and ecosystem services to enhance restoration effectiveness. Proceedings of the National Academy of Sciences. 110 (1): 372-377.

An, K., S. Hwang, D. Bae, J. Lee, and J. Kim. 2005. Introduction of an electrofishing technique for assessments of fish assemblages to Korean watersheds. Korean Journal of Limnology. 38 (4): 482-493.

Argillier, C., S. Caussé, M. Gevrey, S. Pédron, J. De Bortoli, S. Brucet, M. Emmrich, E. Jeppesen, T. Lauridsen, T. Mehner, M. Olin, M. Rask, P. Volta, I. J. Winfield, F. Kelly, T. Krause, A. Palm , and K. Holmgren. 2013. Development of a fish-based index to assess the eutrophication status of European lakes. Hydrobiologia. 704: 193-211.

Baker, E. A., K. E. Wehrly, P. W. Seelbach, L. Wang, M. J. Wiley, and T. Simon. 2005. A multimetric assessment of stream condition in the Northern Lakes and Forests ecoregion using spatially explicit statistical modeling and regional normalization. Transactions of the American Fisheries Society. 134 (3): 697-710.

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

Brenden, T. O., Clark, R. D., Cooper, A. R., Seelbach, P. W., Wang, L., Aichele, S. S. 2006. A GIS framework for collecting, managing, and analyzing multi-scale landscape variables across large regions for river conservation and management. Landscape influences on stream habitats and biological assemblages. American Fisheries Society, Symposium 48. Bethesda, Maryland. pp. 49-74.

Bryce, S. A. and S. E. Clark. 1996. Landscape-level ecological regions: linking state-level ecoregion frameworks with stream habitat classifications, Environmental Management. 20: 297-311.

Cao, Y., C. P. Hawkins, and A. W. Storey. 2005. A method for measuring the comparability of different sampling methods used in biological surveys: implications for data integration and synthesis. Freshwater Biology. 50: 1105-1115.

Clarke, R. T. and D. Hering. 2006. Errors and uncertainty in bioassessment methods – major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia. 566: 433-439.

Davis, L. S. and J. A. Henderson. 1978. Many uses and many users: some desirable characteristics of a common land and water classification system, in classification, inventory, and analysis of fish and wildlife habitat. FWS/OBS-78/76, U. S. Fish and Wildlife Service, Washington, D. C., 13-34.

Diamond, J. M., M. T. Barbour, and J. B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. Journal of North American Benthological Society. 15 (4): 713-727.

ESRI. 2009. ArcGIS. Version 9.1. (Build 722), Redlands, CA.

Furse, M., D. Hering, O. Moog, P. Verdonschot, R. K. Johnson, K. Brabec, K. Gritzalis, A. Buffagni, P. Pinto, N. Friberg, J. Murray-Bligh, J. Kokes, R. Alber, P. Usseglio-Polatera, P. Haase, R. Sweeting, B. Bis, K. Szoszkiewicz, H. Soszka, G. Springe, F. Sporka, and I. Krno. 2006. The STAR project: context, objectives, and approaches. Hydrobiologia. 566: 3-29.

Hawkins, C. P., J. R. Olson, and R. A. Hill. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society 29:312–343.

Hilsenhoff, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20: 31-39.

Houston, L., M. T., Barbour, D. Lenat, and D. Penrose. 2002. A multi-agency comparison of aquatic macroinvertebrate-based stream bioassessment methodologies. Ecological Indicators. 1: 279-292.

Hwang, S., S. Lee, and Y. Park. 2011. Ecological monitoring, assessment, and restoration of running waters in Korea. Annales de Limnologie – International Journal of Limnology. 47: S1-S2.

ITFM (Intergovernmental Task Force on Monitoring Water Quality). 1995. The strategy for improving water – quality monitoring in the United States. Final report of the Intergovernmental Task Force on Monitoring Water Quality. Office of Water Data Coordination, U. S. Geological Survey, Reston, Virginia, Open-File Report 95-742.

Jun, Y., N. Kim, S. Kwon, S. Han, I. Hwang, J. Park, D. Won, M. Byun, H. Kong, J. Lee, and S. Hwang. 2011. Effects of land use on benthic macroinvertebrate communities: Comparison of two mountain streams in Korea. Annales de Limnologie – International Journal of Limnology. 47: S35-S49.

Karr, J. R. 1981. Assessment of biotic integrity using fish communites. Fisheries. 6: 21-27.

Karr, J. R. and E. W. Chu. 2000. Sustaining living rivers. Hydrobiologia. 422: 1-14.

Kilgour, B. W. and L. W. Stanfield. 2006. Hindcasting reference conditions in streams. American Fisheries Society Symposium. 48: 623-639.

KMA (Korea Meteorological Administration). Climate data of South Korea. 2011. Available: Http://www.kma.go.kr. (May 2011).

KOSIS (KOrean Statistical Information Service). 2015. South Korea Summary: 2010 Census data Summary. http://kosis.kr/statisticsList/statisticsList_01List.jsp?vwcd=MT_ZTITLE&parentId=A#SubCont.

Launois, L., J. Veslot, P. Irz, and C. Argillier. 2011. Development of a fish-based index (FBI) of biotic integrity for French lakes using the hindcasting approach. Ecological Indicators. 11: 1572-1583.

Lee, J. H., J. Han, H. K. Kumar, J. Choi, H. K. Byeon, J. Choi, J. Kim, M. Jang, H. Park, and K. An. 2011. National-level integrative ecological health assessments based on the index of biological integrity, water quality, and qualitative habitat evaluation index, in Korean rivers. Annales de Limnologie – International Journal of Limnology. 47: S73-S89.

MEA (Millennium Ecosystem Assessment). 2005. The millennium ecosystem assessment series. Ecosystems and human well-being: Multiscale assessment. Vol. 4. Island Press. NW, Washington DC.

Merritt, R. W., and K. W. Cummins. 1996. An introduction to the aquatic insects of North America. Kendall Hunt Publishing, Dubuque, Iowa.

Michigan Department of Environmental Quality (MDEQ). 1997. Qualtitative biological and habitat survey protocols for wadeable streams and rivers. Michigan Department Environmental Quality (MDEQ), Surface Water Quality Division, Great Lakes and Environmental Assessment Section, GLEAS Procedure 51 (Revised 2002). Lansing, Michigan.

Michigan Department of Environmental Quality (MDEQ). 2002. Procedure 51: Qualitative biological and habitat survey protocols for wadeable streams and rivers. Surface Water Quality Division, Great Lakes and Environmental Assessment Section, GLEAS Procedure 51 (Revised 2002). Lansing, Michigan.

Moulton, S. R., J. G. Kennen, R. M. Goldstein and J. A. Hambrook. 2002. Revised protocols for sampling algal, invertebrate and fish communities as part of the National Water-Quality Assessment Program. USGS Open-File Report 02-150. Reston, Virginia.

NIER (National Institute of Environmental Research). 2009. Investigation and evaluation of aquatic ecosystem of South Korean streams and rivers. Final Report. May 2009. Ministry of Environment. S. Korea.

Park, K. S. 2007. Effects of different sampling effort and taxonomic resolution on assessment metrics. Masters thesis. University of Michigan. Ann Arbor, MI.

Pearl, J. 2009. Causality. Cambridge University Press, New York.

Resh, V. H., and J. K. Jackson. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. pp 195-233. In D. M. Rosenberg and V. H. Resh (eds.). Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Ricklefs, R. E. 2008. The economy of nature. 6th Edition. W. H. Freeman and Co.

Riseng, C. M., M. J. Wiley, R. J. Stevenson, and P. J. Seelbach. 2006. Comparison of coarse versus fine scale sampling on statistical modeling of landscape effects on fish communities of the Muskegon River. Pages 555-576 in R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Landscape influences on stream habitats and biological assemblages. American Fisheries Society, Symposium 48, Bethesda, Maryland.

Riseng, C. M., L. Wang, M. J. Wiley, E. Rutherford, and T. Brendon. 2008. State-of-the-Art approaches for assessment of Great Lakes nearshore and large river habitat. final report to Great Lakes Fishery Trust, U.S. EPA, and Great Lakes Fishery Commission.

Riseng, C. M., M. J. Wiley, P. W. Seelbach, and R. J. Stevensen. 2010. An ecological assessment of Great Lakes tributaries in the Michigan Peninsulas. Journal of Great Lakes Research. 36: 505-519.

Riseng, C. M., M. J. Wiley, R. W. Black, and M. D. Munn. 2011. Impacts of agricultural land use on biological integrity: a causal analysis. Ecological Applications. 21 (8): 3128-3146.

Schoolmaster, D. R. Jr., J. B. Grace, E. W. Schweiger, B. R. Mitchell, and G. R. Guntenspergen. 2013. A causal examination of the effects of confounding factors on multimetirc indices. USGS Saff--Published Research. Paper 769.

Seelbach, P. W., M. J. Wiley, P. A. Soranno, and M. T. Bremigan. 2002. Aquatic conservation planning: using landscape maps to predict ecological reference conditions for specific waters. In: Gutzwiller, K. J. (Editor), Applying Landscape Ecology in Biological Conservation. Springer-Verlag Publishers. New York, New York. pp. 454–478.

SPSS, Inc. 2003. SPSS for Windows Version 12.0. SPSS, Inc., Chicago, Illinois.

Stoddard, J. L., A. T. Herlihy, D. V. Peck, R. M. Hughes, T. R. Whittier, E. Tarquinio. 2008. A process for creating multimetric indices for larger-scale aquatic surveys. Journal of North American Benthological Society. 27: 878-891.

US Census Bureau (U. S. Census Bureau). 2015. United States Summary: 2010 Census Redistricting Data (P. L. 94-171) Summary File. http://www.census.gov/rdo/data/2010_census_redistricting_data_pl_94-171_summary_files.html.

US EPA (U. S. Environmental Protection Agency). 2007. National water quality inventory: report to congress. EPA/841/R-07/001. Office of Water, Washington, DC.

Velleman, P. F. 1997. DataDesk Version 6.0: Statistics Guide. Data Description. Ithaka, NY.

WAMIS (Water Management Information System). 2011. The Ministry of Land, Infrastructure, and Transport. Stream flow gauges and flow data. Republic of Korea. Available: Http://wamis.go.kr.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 2006. Influence of landscape features on summer water temperatures in lower Michigan streams. Pages 113-128 In R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Influences of landscapes on stream habitats and biotic assemblages. American Fisheries Society. Symposium 48. Bethesda, Maryland.

Whittier, T. R., R. M. Hughes, J. L. Stoddard, G. A. Lomincky, D. V., Peck, and A. T. Herlihy. 2007. A structured approach for developing indices of biotic integrity: three examples from streams and rivers in the Western USA. Transactions of American Fisheries Society. 136: 718-735.

Wiley, M. J., P. W. Seelbach, K. Wehrly, and J. S. Martin. 2003. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. pp 201-223. *in*: T.P. Simon, editor. Biological response signatures: indicator patterns using aquatic communities. CRC Press Lewis publishers Inc., Boca Raton, Florida.

Won, D. H., Y. C. Jun, S. J. Kwon, S. J. Hwang, K. G. Ahn, and J. K. Lee. 2006. Development of Korean Saprobic Index using benthic macroinvertebrates and its application to biological stream environment assessment. Journal of Korean Society on Water Quality. 22: 768-783.

Yoon, J, J. Kim, M. Byeon, H. Yang, J. Park, J. Shim, H. Song, H. Yang, and M. Jang. 2011. Distribution patterns of fish communities with respect to environmental gradients in Korean streams. Annales de Limnologie – International Journal of Limnology. 47: S63-S71.

Zelinka, M. and P. Marvan. 1961. Zur präzisierung der biologische klassifikation der reinheit fliessender gewässer. Archiv fur Hydrobiologie. 57: 389-407.

Zonneveld, I. S. 1994. Basic principles of classification, in F. Klijn (Editor). Ecosystem classification for environmental management, Kluwer Academic Publishers, Dordrecht, The Netherlands, 23-47.

Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2002. Distributions of stream fishes and their relationship to stream size and hydrology in Michigan's Lower Peninsula, Transactions of the American Fisheries Society. 131: 70-85.