

工學博士學位請求論文

오염하천 생태계 관리를 위한 생태정보학 기반  
통합수질지수 연구

*Study of Integrative Water Quality Index Based on  
Ecological Informatics for Management of the Polluted  
Stream Ecosystems*

指導教授 高星澈

2008年 2月

韓國海洋大學校 大學院

土木環境工學科 環境工學專攻

崔正惠

本 論 文 을 崔 正 惠 의 工 學 博 士 學 位 論 文 으 로 認 准 함.

委 員 長 工 學 博 士 金 仁 洙 (印)

委 員 理 學 博 士 高 星 澈 (印)

委 員 工 學 博 士 宋 永 彩 (印)

委 員 理 學 博 士 金 明 珍 (印)

委 員 理 學 博 士 全 胎 秀 (印)

2008年 2月

韓 國 海 洋 大 學 校 大 學 院  
土 木 環 境 工 學 科 環 境 工 學 專 攻  
崔 正 惠

# List of Contents

|  |      |
|--|------|
| Table of Contents .....                      | i    |
| List of Figures .....                        | v    |
| List of Tables .....                         | xi   |
| Abstract .....                               | xiii |
| <br>   |      |
| 제 1장 서론 .....                                | 1    |
| <br>   |      |
| 제 2장 문헌연구 .....                              | 4    |
| 2.1. 수질평가 방법 .....                           | 4    |
| 2.1.1. 화학적 수질평가 .....                        | 4    |
| 2.1.2. 생물학적 수질평가 .....                       | 6    |
| 2.1.3. 통합적 수질평가 .....                        | 9    |
| 2.2. 수질평가 현황 .....                           | 11   |
| 2.2.1. 국외 수질평가 현황 .....                      | 11   |
| 2.2.2. 국내 수질평가 현황 .....                      | 13   |
| 2.2.2.1. 국내의 수질 지표 개발 현황 .....               | 16   |
| 2.2.2.2. 국내 수질환경관리를 위한 통합적 수질지표개발의 필요성 ..... | 18   |
| 2.2.2.3. 통합 수질지표 개발 방향 .....                 | 19   |
| <br>   |      |
| 제 3장 미소 생태계 환경 및 생물요인 .....                  | 21   |
| 3.1. 실험방법 .....                              | 21   |
| 3.1.1. 조사지점 .....                            | 21   |
| 3.1.2. 미소생태계의 이화학적 환경요인 .....                | 24   |
| 3.1.3. 미소 생태계의 군집구조 .....                    | 25   |
| 3.1.3.1. 저서성 대형무척추동물 군집구조 .....              | 25   |
| 3.1.3.2. 부착조류 및 부유조류 군집구조 .....              | 25   |
| 3.1.3.3. 미생물 군집구조 .....                      | 26   |

|   |    |
|---|----|
| 3.2. 결과 및 고찰 .....                                    | 28 |
| 3.2.1 미소생태계의 이화학적 환경요인 .....                          | 28 |
| 3.2.1.1. 수온, 유속 및 수심 .....                            | 28 |
| 3.2.1.2. pH .....                                     | 31 |
| 3.2.1.3. Conductivity .....                           | 32 |
| 3.2.1.4. Turbidity .....                              | 33 |
| 3.2.1.5. Biochemical oxygen demand .....              | 34 |
| 3.2.1.6. Total organic carbon .....                   | 35 |
| 3.2.1.7. Total phosphorus .....                       | 36 |
| 3.2.1.8. Total nitrogen .....                         | 37 |
| 3.2.2 미소생태계의 군집구조 .....                               | 38 |
| 3.2.2.1. 저서성 대형 무척추동물 군집구조 .....                      | 38 |
| 3.2.2.2. 부착 및 부유 조류 군집구조 .....                        | 40 |
| 3.2.2.3. 미생물 군집구조 .....                               | 43 |
| 3.3. 결론 .....   | 48 |
| <br>  |    |
| 제 4장 통합지수개발을 위한 사전 연구 .....                           | 49 |
| 4.1. 연구배경 .....                                       | 50 |
| 4.2. 실험방법 .....                                       | 51 |
| 4.2.1. Self-organizing mapping(SOM)을 이용한 군집 유형화 ..... | 51 |
| 4.2.2. 다층퍼셉트론을 이용한 생물군집과 환경요인의 연관성 .....              | 54 |
| 4.2.3. 생물 지수 .....                                    | 56 |
| 4.3. 결과 및 고찰 .....                                    | 57 |
| 4.3.1. Self-organizing mapping(SOM)을 이용한 군집 유형화 ..... | 57 |
| 4.3.1.1. 미생물 .....                                    | 57 |
| 4.3.1.2. 대형무척추동물 .....                                | 63 |
| 4.3.1.3. 부착조류 .....                                   | 69 |
| 4.3.2. 다층퍼셉트론을 이용한 생물군집과 환경요인의 연관성 .....              | 73 |
| 4.3.3. 다분류군 생물지수에 대한 환경요인과의 관계 .....                  | 75 |
| 4.3.4. 각 분류군 생물지수의 연관성 .....                          | 78 |
| 4.3.5. 다분류군의 군집지수 연관성 .....                           | 81 |

|   |     |
|---|-----|
| 4.4. 결론 .....                               | 83  |
| <br>  |     |
| 제 5장 통합지수 개발을 위한 미생물 지수 개발 .....            | 84  |
| 5.1. 연구배경 .....                             | 85  |
| 5.2. 실험방법 .....                             | 86  |
| 5.2.1. 데이터 분석전략 .....                       | 86  |
| 5.2.2. 요인분석에 의한 환경요인의 추출 및 환경변수와의 관계 .....  | 88  |
| 5.2.3. 환경요인에 대한 확률모델 및 오염지역 분류 .....        | 91  |
| 5.2.4. 확률모델을 이용한 미생물 출현패턴 분석 .....          | 94  |
| 5.2.5. 상대적 중요도 평가를 통한 주요 미생물 추출 .....       | 97  |
| 5.2.6. 미생물 지수화 .....                        | 102 |
| 5.3. 결과 및 고찰 .....                          | 104 |
| 5.3.1. 요인분석에 의한 환경요인의 추출 및 환경변수와의 관계 .....  | 104 |
| 5.3.2. 요인분석에 의해 추출된 환경요인과 미생물 군집과의 관계 ..... | 112 |
| 5.3.3. 환경요인에 대한 확률모델 및 오염지역 분류 .....        | 115 |
| 5.3.4. 확률모델을 이용한 미생물 출현패턴 분석 .....          | 121 |
| 5.3.5. 상대적 중요도 평가를 통한 주요 미생물 추출 .....       | 127 |
| 5.3.6. 미생물 지수화 및 환경평가 .....                 | 133 |
| 5.4. 결론 .....                               | 140 |
| <br>  |     |
| 제 6장 대형무척추동물 지수 및 환경평가를 통한 유효성 검토 .....     | 141 |
| 6.1. 실험방법 .....                             | 143 |
| 6.1.1. 대형무척추동물 지수화 .....                    | 143 |
| 6.1.2. 지수의 유효성 검토 .....                     | 144 |
| 6.2. 결과 및 고찰 .....                          | 145 |
| 6.2.1. 확률모델을 이용한 대형무척추동물 출현패턴 분석 .....      | 145 |
| 6.2.2. 상대적 중요도 평가를 통한 주요 대형무척추동물 추출 .....   | 150 |
| 6.2.3. 대형무척추동물 지수화 및 환경평가 .....             | 155 |
| 6.2.4. 유효성 검증 .....                         | 160 |
| 6.3. 결론 .....                               | 161 |

|                                  |     |
|----------------------------------|-----|
| 제 7장 통합지수 개발 .....               | 162 |
| 7.1. 실험방법 .....                  | 163 |
| 7.2. 결과 및 고찰 .....               | 164 |
| 7.2.1. 미생물-대형무척추동물 통합지수 .....    | 164 |
| 7.2.2. 미생물-대형무척추동물-조류 통합지수 ..... | 167 |
| 7.2.3. 통합지수 평가 .....             | 170 |
| 7.4. 결론 .....                    | 173 |
| <br>                             |     |
| 제 8장 종합결론 .....                  | 174 |
| <br>                             |     |
| 제 9장 참고문헌 .....                  | 176 |

# *List of Figure*

|   |    |
|---|----|
| Fig. 3.1. Map of the sampling sites for this study .....  | 22 |
| Fig. 3.2. Procedure for analysis of microbial communities using<br>PCR-DGGE .....   | 27 |
| Fig. 3.3. Average temperature of water samples from the sampling sites<br>over the entire sampling period .....               | 29 |
| Fig. 3.4. Average velocity of stream water from the sampling sites over the<br>entire sampling period .....                   | 29 |
| Fig. 3.5. Average depth of stream water from the sampling sites over the<br>entire sampling period .....                      | 30 |
| Fig. 3.6. Average pH of water samples from the sampling sites over the<br>entire sampling period .....                        | 31 |
| Fig. 3.7. Average conductivity of water samples from the sampling sites<br>over the entire sampling period .....              | 32 |
| Fig. 3.8. Average turbidity of water samples from the sampling sites over<br>the entire sampling period .....                 | 33 |
| Fig. 3.9. Average biochemical oxygen demand of water samples from the<br>sampling sites over the entire sampling period ..... | 34 |
| Fig. 3.10. Average total organic carbon of water samples from the sampling<br>sites over the entire sampling period .....     | 35 |
| Fig. 3.11. Average total phosphorus of water samples from the sampling<br>sites over the entire sampling period .....         | 36 |
| Fig. 3.12. Average total nitrogen of water samples from the sampling sites<br>over the entire sampling period .....           | 37 |
| Fig. 3.13. Relative abundance of the macroinvertebrate community from the<br>pollution sources .....                          | 39 |
| Fig. 3.14. Relative abundance of the attached algae community from the<br>pollution sources .....                             | 41 |
| Fig. 3.15. Number of the attached algae species from the<br>pollution sources .....   | 41 |

|  |    |
|--|----|
| Fig. 3.16. PCR-DGGE profile of microbial communities .....   | 41 |
| Fig. 3.17. Cluster analysis of microbial communities based on species<br>richness over the entire sampling period .....  | 45 |
| Fig. 4.1. Schematic diagram of the Kohonen network .....   | 51 |
| Fig. 4.2. SOM patterning of sampling sites based upon microbial density.<br>Sites: Duk, Ddk, Dks, Dkh, Dag .....   | 58 |
| Fig. 4.3. Distribution of microbial populations on the SOM. Sites: Duk, Ddk,<br>Dks, Dkh, Dag .....  | 59 |
| Fig. 4.4. SOM patterning of sampling sites based upon microbial density.<br>Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD,<br>InJ .....                     | 60 |
| Fig. 4.5. Distribution of microbial populations on the SOM. Sites: ReA, ReT,<br>ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ .....                                     | 61 |
| Fig. 4.6. SOM patterning of sampling sites based upon microbial density.<br>Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD .....  | 62 |
| Fig. 4.7. Distribution of microbial populations on the SOM. Sites: DUK,<br>DDK, DKS, DAG, OCU, ONS, SSI, HJD .....   | 62 |
| Fig. 4.8. SOM patterning of sampling sites based upon macroinvertebrate<br>population density. Sites: Duk, Ddk, Dks, Dkh, Dag .....  | 64 |
| Fig. 4.9. Distribution of macroinvertebrate population on the SOM. Sites:<br>Duk, Ddk, Dks, Dkh, Dag .....   | 65 |
| Fig. 4.10. SOM patterning of sampling sites based upon macroinvertebrate<br>population density. Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC,<br>FaA, LiT, LiK, HJD, InJ ..... | 66 |
| Fig. 4.11. Distribution of macroinvertebrate population on the SOM. Sites:<br>ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK,<br>HJD, InJ .....                          | 67 |
| Fig. 4.12. SOM patterning of sampling sites based upon macroinvertebrate<br>population density. Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI,<br>HJD .....                     | 68 |



|   |     |
|---|-----|
| Fig. 4.13. Distribution of macroinvertebrate population on the SOM. Sites:<br>DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD .....  | 68  |
| Fig. 4.14. SOM patterning of sampling sites based upon attached algal<br>population density. Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC,<br>FaA, LiT, LiK, HJD, InJ ..... | 70  |
| Fig. 4.15. Distribution of attached algal population on the SOM. Sites: ReA,<br>ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ .....                             | 71  |
| Fig. 4.16. SOM patterning of sampling sites based upon attached algal<br>population density. Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI,<br>HJ .....                      | 72  |
| Fig. 4.17. Distribution of attached algal population on the SOM. Sites: DUK,<br>DDK, DKS, DAG, OCU, ONS, SSI, HJD .....   | 72  |
| Fig. 4.18. Recognition of species richness using multi-layer perceptron .....   | 73  |
| Fig. 4.19. Sensitivity analysis of species richness by MLP using BOD and<br>DO as input data .....  | 74  |
| Fig. 4.20. Effect of COD on species richness of the inter-taxa .....  | 76  |
| Fig. 4.21. Effect of BOD on species richness of the inter-taxa .....  | 77  |
| Fig. 4.22. Grouping of indices for the inter-taxa by SOM .....  | 79  |
| Fig. 4.23. Distribution of each taxon depending on species diversity .....  | 82  |
| Fig. 4.24. Distribution of each taxon depending on species richness .....   | 82  |
| Fig. 5.1. Dimensionality reduction and principal-factor extraction through<br>factor analysis technique .....   | 88  |
| Fig. 5.2. Schematic diagram probabilistic neural networks .....   | 98  |
| Fig. 5.3. Concept of reject boundary in input space using a radial basis<br>function with gaussian kernel .....   | 98  |
| Fig. 5.4. Structure of the proposed system for measuring the relative<br>importance among microorganisms .....  | 100 |
| Fig. 5.5. A result of mapping the environment variables into principal factor<br>space .....  | 105 |

|  |     |
|--|-----|
| Fig. 5.6 A comparative analysis to reveal the relation between common factor and turbidity variable through regression and factor analysis .....                         | 109 |
| Fig. 5.7 A comparative analysis to reveal the relation between common factor and BOD variable through regression and factor analysis ..                                  | 109 |
| Fig. 5.8 A comparative analysis to reveal the relation between common factor and T-P variable through regression and factor analysis ..                                  | 110 |
| Fig. 5.9 A comparative analysis to reveal the relation between common factor and T-N variable through regression and factor analysis ..                                  | 110 |
| Fig. 5.10 A comparative analysis to reveal the relations between common factor and NO <sub>3</sub> <sup>-</sup> -N variable through regression and factor analysis ..... | 111 |
| Fig. 5.11. Correlation between the common factor and <i>Acinetobacter</i> sp. ..   | 113 |
| Fig. 5.12. Correlation between the common factor and <i>Sphingomonas</i> sp. ..  | 113 |
| Fig. 5.13. Correlation between the common factor and <i>Pseudomonas</i> sp. ..   | 114 |
| Fig. 5.14. Probabilistic model with the two gaussian distribution functions for the factor scores .....  | 119 |
| Fig. 5.15. Two-categories classification based on Bayesian decision rule for the factor scores .....   | 120 |
| Fig. 5.16. Probability distribution in terms of characterizing microbial community pattern as P1 .....   | 123 |
| Fig. 5.17. Probability distribution in terms of characterizing microbial community pattern as P2 .....   | 124 |
| Fig. 5.18. Probability distribution in terms of characterizing microbial community pattern as P3 .....   | 125 |
| Fig. 5.19. Probability distribution in terms of characterizing microbial community pattern as P4 .....   | 126 |
| Fig. 5.20. Relative importance of the indicators selected by PNN-based evaluation system .....   | 129 |
| Fig. 5.21. Relative importance of the microbial populations as an indicator in relatively clean site .....   | 130 |

|   |     |
|---|-----|
| Fig. 5.22. Factor–score distributions for the microbial populations as an indicator in relatively clean site .....  | 130 |
| Fig. 5.23. Relative importance of the microbial populations as an indicator in relatively polluted site .....   | 131 |
| Fig. 5.24. Factor–score distributions for the microbial populations as an indicator in relatively polluted site .....   | 131 |
| Fig. 5.25. Factor–score distributions for the microbial populations as a indicator in both clean and polluted site .....  | 132 |
| Fig. 5.26. Joint distribution and correlation between the factor and <i>Acinetobacter</i> sp. as a predictor in the pattern of P4 .....                                     | 132 |
| Fig. 5.27. Index scores of the microbial populations selected as an indicator in relatively polluted and clean site as a result of the relative importance evaluation ..... | 134 |
| Fig. 5.28. An assessment of water quality using microbial indices .....   | 138 |
| Fig. 6.1. Probability distribution in terms of characterizing macroinvertebrate community pattern as P1 .....   | 146 |
| Fig. 6.2. Probability distribution in terms of characterizing macroinvertebrate community pattern as P2 .....   | 147 |
| Fig. 6.3. Probability distribution in terms of characterizing macroinvertebrate community pattern as P3 .....   | 148 |
| Fig. 6.4. Probability distribution in terms of characterizing macroinvertebrate community pattern as P4 .....   | 149 |
| Fig. 6.5. Relative importance of the indicators selected by PNN–based evaluation system .....   | 151 |
| Fig. 6.6. Relative importance of the macroinvertebrate populations as an indicator in relatively clean site .....   | 152 |
| Fig. 6.7. Factor–score distributions for the macroinvertebrate populations as an indicator in relatively clean site .....   | 152 |
| Fig. 6.8. Relative importance of the macroinvertebrate populations as an indicator in relatively polluted site .....  | 153 |

|   |     |
|---|-----|
| Fig. 6.9. Factor–score distributions for the macroinvertebrate populations as an indicator in relatively polluted site .....  | 153 |
| Fig. 6.10. Factor–score distributions for the macroinvertebrate populations as a indicator in both clean and polluted site .....  | 154 |
| Fig. 6.11. Index scores of the macroinvertebrate populations selected as an indicator in relatively polluted and clean site as a result of the relative importance evaluation ..... | 156 |
| Fig. 6.12. An assessment of water quality using macroinvertebrate indices   | 159 |
| Fig. 6.13. Efficacy assessment of the developed macroinvertebrate index in comparison with the BMWP index .....   | 160 |
| Fig. 7.1. An assessment of water quality using the integrative water quality index based upon the multi–taxa (microorganisms and macroinvertebrates) .....                          | 166 |
| Fig. 7.2. An assessment of water quality using the integrative water quality index based upon the multi–taxa (microorganisms, macroinvertebrates and algae) .....                   | 169 |
| Fig. 7.3. Comparison of the estimation performance for water quality assessment in four cases of integrative water quality indices .....  | 172 |

# *List of Table*

|  |     |
|--|-----|
| Table 2.1. Environmental regulation criteria for water quality of the streams in South Korea .....                               | 14  |
| Table 2.2. Biological characteristics across the levels of water quality in South Korea .....                                    | 15  |
| Table 3.1. Sampling sites for this study .....   | 23  |
| Table 3.2. Community composition of macroinvertebrates depending on the pollution sources .....                                  | 39  |
| Table 3.3. Biotic indices of the attached algae depending on pollution source .....  | 42  |
| Table 3.4. Identification of selected PCR-DGGE bands based on 16S rDNA sequence .....  | 46  |
| Table 4.1. Pearson correlation coefficients for indices of the inter-taxa .....  | 79  |
| Table 4.2. Pearson correlation coefficients for environmental parameters and biological indices .....                            | 80  |
| Table 5.1. Factor loadings of 4 principal factors extracted by factor analysis for all of the environment variables tested ..... | 105 |
| Table 5.2. Factor loadings for the 5 environment variables selected by factor analysis .....                                     | 106 |
| Table 5.3. Coefficients for mapping factor-space into variable-space for the selected environment variables .....                | 107 |
| Table 5.4. Two-category classification based on probabilistic model and average scores of factors in the sampling sites .....    | 116 |
| Table 5.5. Indices and scores of the selected microbial populations .....  | 135 |
| Table 5.6. Numerical relation between microbial index and each of the environment variables .....                                | 136 |
| Table 5.7. Comparison of average index scores with average factor scores for the sampling sites .....                            | 139 |

|  |     |
|--|-----|
| Table 6.1. Index and the score of selected macroinvertebrate populations   | 157 |
| Table 6.2. Comparison of average index scores with average factor scores<br>for the sampling sites                     | 158 |
| Table 7.1. Scores for the integrated indices based upon the multi-taxa<br>(microorganisms and macroinvertebrate)       | 165 |
| Table 7.2. Score for the integrated index based upon multi-taxa<br>(microorganism, macroinvertebrate and algae)        | 168 |
| Table 7.3. Comparison of estimation performance for each biological index in<br>terms of IAE (Integral Absolute Error) | 171 |

## *Abstract*

A significant amount of various wastewaters including domestic and industrial wastewaters was increasingly discharged into the environment since 1980s in South Korea which experienced a rapid population growth and a fast industrialization and urbanization, causing a serious ecological perturbation and destruction in aquatic ecosystems. Physico-chemical parameters have been conventionally used to monitor water quality, and then biological indicators were introduced to monitor water qualities in the late 1970s for streams, lakes and other water bodies (Wang, 2001). One of the major limitations in management of water quality in the field is that the management is performed based on maximum allowable concentration levels for control parameters, and a number of biological indicators which reflect a limited information of biological influences caused by the pollution.

Physico-chemical monitoring of aquatic environments is effective in providing a rapid and specific information for water quality management but it has a significant limitation in demonstrating an overall status of an aquatic ecosystem affected by the exposure of biological communities to the various pollutants. Therefore, it will be necessary to develop an integrative water quality index which makes possible a comprehensive analysis of water quality based on an understanding of ecosystem function on spatial and temporal scales.

The purpose of this study was to develop an integrative water quality index for the management of stream ecosystems. In order to effectively evaluate the water quality for the ecosystems, it will be necessary to combine physicochemical parameters and biological parameters complementarily, which may comprehensively represent the quality of the ecosystem. This study, therefore, focused on development of an integrative water quality index that took into account the community structures of producers (algae), consumers (macro-invertebrates) and degraders (microorganisms) and their environmental factors as the ecosystem

components.

Data for the biological and environmental parameters were periodically collected from 21 sites in the five different pollution sources (domestic sewage, farming, livestock, industry and restoration sites). The biological parameters included population densities of algae, macroinvertebrates and microorganisms, and the environmental parameters were atmospheric temperature, water temperature, pH, conductivity, DO, BOD, turbidity, water depth, flow rate, COD, TOC, T-P, T-N and  $\text{NO}_3^-$ -N. The microbial community analysis was performed by PCR-DGGE (denaturing gradient gel electrophoresis) of amplified 16S rDNA fragments. Total data collected were 3,289 (13 items  $\times$  253 sites) for the environmental parameters, 21,505 (85 family  $\times$  253 sites) for macro-invertebrates, 29,095 (115 species  $\times$  253 sites) for microorganisms and 15,180 (60 species  $\times$  253 sites) for algae.

All the environmental and biological parameters were first processed through SOM analysis to pattern the complex ecological data sets. SOM patterning of sampling sites using macroinvertebrate densities as an input showed generally a good correlation between the site pollution status and their relevant species while there was little correlation in case of microorganisms and algae. In fact, the macroinvertebrate community was clustered according to the pollution gradient while microbial and algal communities appeared to cluster on a seasonal basis.

The SOM analysis was also performed for the representative indices calculated from these taxa data. This procedure was necessary to obtain an essential data useful for development of the integrative water quality index. The correlations between macroinvertebrate and algal communities were significant to a certain level while the microbial community showed little correlations with the other taxa. The correlation analyses between environmental parameters and the three taxa indicated that each taxon reflected the water quality independently. Microbial index has been barely developed and the microbial populations showed little correlations with the environmental parameters in this study, which made difficult the



development of the index.

To circumvent this problem, an attempt has been made to develop a microbial water quality index that could represent the environmental status. For development of the index, microbial populations which could well represent the environmental status were first selected and characterized based upon multi-variate analysis (factor analysis) and the probability model of the microbial communities. This trial, however, was not successful because of non-linear responses of the microbial populations to the environmental parameters.

The factor analysis made it possible to extract a common factor out of the high-dimensional environment variables (parameters) that could well represent a pollution status of the environment. The data for the common factor were essentially classified into clean site ( $w_1$ ) and relatively polluted site ( $w_2$ ), leading to a better elucidation of relationships between the common factor and the environmental parameters. Emergence patterns for the microbial populations were analyzed based on the probabilistic model derived from the microbial densities for each site. The microbial index for the water quality was developed by utilizing characters and emergence pattern of each microbial population. A water quality index based on macroinvertebrate community data was also developed in the same way. Each index developed from microbial and macroinvertebrate community data was evaluated against the measured environmental data to see if they could reflect the environmental parameters well. The evaluation results for the microbial and macroinvertebrate indices generally showed a good predictability for the water quality. However, their prediction capability becomes lower in the highly polluted area (*e.g.*, industrial pollution sites) and the mid-level pollution (BOD, 14.65 ~ 35.08 mg/L ) sites. By the way, the microbial index was able to better predict the water quality in the extremely polluted (BOD, 44.66 ~ 76 mg/L) environment than the macroinvertebrate index. This is mainly because most of the microbial population data were collected from the clean sites and the highly polluted

sites, and hence the data were less representative of mid-level pollution sites, and had more missing values than the data of environmental parameters and the other taxa. To avoid this kind of bias, it will be necessary to collect data from more diverse sites of different levels of pollution and to secure more diversified microbial populations that could represent the environmental pollution status more extensively. Furthermore, the macroinvertebrates cannot be a good indicator for the extremely polluted sites and oligotrophic sites since they could not survive at these sites.

The developed macroinvertebrate index was also compared with the indices of BMWP for macroinvertebrates to test and evaluate the efficacy of the index. It was shown that the site evaluation results based upon the indices developed in this study were similar to those based upon BMWP. It was concluded that the integrative index developed in this study was able to predict the water quality in the polluted streams when compared with the conventional BMWP index of macroinvertebrates used to monitor the water quality of the stream ecosystems.

## 제 1장 서론

1980년대 이후 진행된 급격한 인구증가와 산업화 및 도시화로 인해 생활오수 및 각종 폐수의 양이 날로 늘어나고 있어 생태계 교란 현상이 급격히 진행되고 있으며, 이로 인한 수 환경 파괴가 급속히 진행되고 있다. 이에 수질관리의 일환으로 수질을 평가하기 위한 화학적, 물리적 및 생물학적 지표와 같은 다양한 수질평가 방법들이 개발되었다. 최근까지는 물리·화학적 지표를 기반으로 하는 수질평가법이 주로 이용되었고, 생물학적 지표를 이용한 평가방법이 1970년대 말부터 1980년대 초에 개발된 이래 하천, 호수 또는 기타 수계의 생태계 건강도나 상태를 측정하는데 있어 성공적으로 적용되어 오고 있다(Wang, 2001).

실제 환경관리 현장에서는 환경관련법에 나타나 있는 기준에 의거하여 일정한 기준을 넘지 않도록 규제하는 방향으로 시행되고 있고, 인간의 위생적인 측면만 강조되고 생태계 대사나 전반적인 추이는 고려하지 않고 있다. 관리기준도 pH, DO, BOD, 부유물질 및 중금속 함유도 등과 같은 특정한 측정목표 하에 일부 단일 항목이나 수계의 유리한 항목으로 평가해 옴으로서 수질의 객관적인 평가가 어려운 점이 있다. 이화학적 방법은 비교적 빠른 시간에 정밀하고 국소적인 정보를 알려줄 수는 있지만 순간적, 단편적이어서 생태계의 전체적인 상황을 알려주는데 제한이 있다. 이화학적 지표(BOD 또는 COD) 위주의 수질평가는 인체와 생물, 생태계 미치는 영향을 이들 지표가 제대로 반영하지 못하는 한계가 있다.

이러한 단점을 보완하기 위해 1981년 미국 환경청에서는 수자원 평가를 위한 도구로서 어류를 이용한 IBI(index of biotic integrity) 개념을 도입하였다. IBI는 미국 이외에도 프랑스, 아프리카 기니아, 나미비아, 인도, 남미 베네주엘라, 호주 및 일본 등에서 활발히 진행되고 있다(염 등, 2000). 국내의 경우도 어류 등 생물중에 의한 생물학적 평가를 포함하는 수질환경기준이 2008년부터 법제화 될 전망이다. 환경부에 따르면 현재 생물화학적 산소요구량(BOD)과 화학적 산소요구량(COD) 등으로 대표되는 수질환경기준이 2007년부터 국민건강과 생태계 관리 중심으로 바뀌게 될 예정이다. 즉 안전하고 생태적으로 건강한 물 환경 조성을 위해 현재 9개 항목에 불과한 수질환경기준을 2015년까지

30개가량으로 늘리기로 하고 관계법령을 개정, 2007년부터 단계적으로 시행하기로 하였다.

생물 지표를 이용한 생물지수는 생태계 질을 종합적, 지속적으로 표출해 줄 수 있다. 나아가 군집의 종이 다양하므로 군집의 전체적인 변이를 통해 생태계 기능이나 구조에 대한 환경 영향을 나타내거나, 각 종별로 특정 오염원을 검출할 수 있는 장점이 있다. 따라서 오염 하천 생태계의 주요 군집인 생산자(조류), 소비자(무척추동물) 분해자(미생물)의 에너지 흐름에 있어서의 분해능을 파악하는 것이 생태적 수질관리를 위해서 필수적이다. 또한 관리지표에 생물, 생태적 방법을 포함하고 이화학적 지표와 상보적으로 적용하여야 효과적인 수질의 진단이 가능하다.

생물지표를 이용한 평가방법은 어류를 포함한 수서곤충(insects), 대형무척추동물(macroinvertebrates) 그리고 조류(algae) 등과 같은 다양한 생물종을 이용한 평가가 이루어지고 있다. 하지만, 국내외의 연구동향은 어류 및 대형 생물에 국한되어있다(이 등, 2004; 김과 안, 2004; Roth et al., 1998; Bozzetty et al., 2004; Lafontaine et al., 2000). 어류보다 민감도 및 특이성에서 우수한 미생물을 이용한 통합적 수질지표개발연구는 초기 단계이며(장과 이, 2003; 최와 신, 1997; 김 등, 2002) 미생물을 이용한 오염수질환경의 모니터링을 위한 통합수질 지표개발은 거의 전무한 실정이다. 또한 현재까지 수질상태의 평가는 이화학적 지표나 생물지수(조류, 무척추동물 및 미생물) 중 중요 분류군에 대한 별도의 연구가 진행되어져왔다.

기존에 연구가 많이 수행되어진 저서성 대형무척추동물은 종풍부도(SR), 총밀도(TD), EPT 풍부도(EPT-R), EPT 밀도(EPT-D), BMWP 지수((biological monitoring working party: BMWP), 다양도지수(H) 등이, 부착조류의 경우 종풍부도(SR), 다양도지수(H), 총밀도(TD) 및 DAIPo(diatom assemblage index of organic water pollution) 등 다양한 지수들이 개발되었다.

소비자인 저서성 대형무척추동물 군집은 BMWP(biological monitoring working party)를, 생산자인 조류 군집은 DAIPo(diatom assemblage index of organic water pollution)등을 통해 오염에 따른 분류군의 출현 유효 평가가 이루어지고 있다. 그러나 미생물의 경우 미생물과 환경간의 관계를 직접적으로 대표할 수 있는 지수가 전무하다. 또한 대형무척추동물 및 조류 등의 생물지표를 이용한 수질평가방법은 오염 정도를 반영해주는 장점이 있으나 일정

수준이상으로 오염이 심해지면 대상 생물종의 생존이 어렵다. 오염이 심한 수계에서 관찰 대상종이 관찰되지 않음으로서 해당 수계의 환경을 평가하기가 어렵다는 한계가 있다. 그러나 미생물의 경우 오염정도에 관계없이 일정수준의 개체수를 유지하며, 오염에 따른 군집변화를 타나낼 수 있으므로 대형무척추동물 및 조류 등으로 평가할 수 없는 수계의 수질을 평가하는데 대안이 될 수 있다.

이에 본 연구에서는 하천생태계 관리를 위한 미생물 지수를 개발하고자 하였다. 나아가 생태계의 구성요소인 생산자(부착조류), 소비자(대형저서생물) 및 분해자(미생물)의 군집구조와 환경요인(BOD, T-N, T-P 및 TOC 등)을 각각 분석한 후 이들의 자료를 통합적으로 분석하여 지수를 도출하고자 하였다. 분해자인 미생물군집 구조는 분자생물학적기법(PCR-DGGE)을 이용하여 분석하였고 주요 환경요인으로는 수온, pH, conductivity, DO, BOD, 탁도, 수심, 유속, COD, TOC, T-P, T-N 및  $\text{NO}_3^-$ -N 등을 분석하였다. 환경요인과 미생물 군집구조 분석을 통해 수집된 데이터는 확률모델을 기반으로 오염도에 따른 출현 미생물 군집 패턴을 분석하고 환경영향을 잘 반영할 수 있는 미생물을 선별·특성화하여 미생물 지수를 개발하고자 하였다. 이를 통해 개발된 미생물 지수는 타 분류군과 종합적으로 분석하여 통합수질지수를 개발하고자 하였다. 제시된 통합수질지수는 생태계의 오염도를 판정할 뿐 아니라 생태계의 변화과정(천이)을 추적하는데 활용할 수 있을 것으로 판단된다.

## 제 2장 문헌연구

### 2.1. 수질평가 방법

#### 2.1.1. 화학적 수질평가

화학적 수질평가 방법은 화학적 수질지표를 이용하는 방법으로 현재까지 가장 활발히 이용되고 있다. 일반적으로 사용되고 있는 화학적 수질지표는 pH, 전기전도도(conductivity), 용존산소(dissolved oxygen: DO), 질산성 질소(nitrate-N), ortho-phosphates, 화학적산소요구량(chemical oxygen demand: COD), 생화학적산소요구량(biochemical oxygen demand: BOD) 및 pesticides 등이 있다. 그 외에도 온도(temperature), 총 부유물질(total suspended solids: TSS) 및 탁도(turbidity) 등과 같은 물리적 요인도 수질평가를 위해 이용되고 있다. 현재까지 개발된 화학적 수질평가 방법은 위와 같은 물리, 화학적 수질지표의 측정치를 그대로 이용한 단일항목지표가 주로 이용되었고, 1965년 Horton's index를 시작으로 지표항목 측정치를 조합하여 점수로 표시하는 수질지수(WQI: water quality index)들이 개발되었다.

WQI는 다양한 수계의 수질을 비교하기 위한 표준화 방법으로 1970년 Brown, McClelland, Deininger 및 Tozer 등에 의해 제시되었다. 이 연구는 미국 국가위생국(national sanitation foundation: NSF)에 의해 지원되었기 때문에 NSF-WQI로 알려져 있다. NSF-WQI는 미국전역의 수질전문가 상당수의 의견을 조합하여 주관적인 판단을 최소화하려고 시도하였다(Brown, et al., 1970). NSF-WQI는 온도, pH, DO, 탁도, 분변성 대장균군, BOD, T-P, 질산성 질소 및 총 부유물질 등 총 9개의 수질항목의 측정결과에 대해 각 수질항목들은 정해진 환산표에 의해 Q-value로 변환된다. 변환된 각 항목의 Q-value는 상대적 가중치를 고려하여 WQI score로 환산되며 그 범위는 0(오염) ~ 100(청정)이다.

그 외에 Deiniger (1971) 등은 수온, pH, EC, BOD, DO, DSS, NO<sub>3</sub>-N,

turbidity, Phenols, FC, Chromaticity 및 Hardness 등 11개 항목에 대해 가중치를 적용한 WQI를 제시하였고(Deiniger, et al., 1971), 국내에서도 pH, DO, BOD, SS 및 총대장균군 등 5개 항목에 대한 평점을 합하여 항목수로 나누어 하천 오염도를 평가하는 KOE-WQI(Korea Office of Environment Water Quality Index) 등이 개발되었다(환경부, 1989).

단일항목을 이용하여 수질을 평가하는 경우 수질상황을 충분히 반영하지 못하는 단점이 있으나 적용성의 용이함으로 인해 아직까지 상당한 국가들에서 공식적인 수질평가 방법으로 사용되고 있다. 단일항목 측정을 통한 수질평가법이 가지는 단점을 피하기 위해 개발된 WQI의 경우 다양한 환경에 적용할 수 있는 수질지표를 개발하기 위해 많은 연구가 수행되어 지고 있다.

## 2.1.2. 생물학적 수질평가

생물학적 수질평가 방법은 생물지표를 이용하는 방법으로 1800년대 후반 수중생물의 수와 물의 청정도를 연결하는 연구를 수행함으로써 시작되었다. 그러나 1900년대 물리-화학적 분석기술의 발달로 크게 각광받지 못하다가 1980년 이후 미국을 중심으로 다시 활발히 연구되고 있다. 생물학적 평가방법은 오염을 종합적으로 평가할 수 있는 장점을 가지고 있다. 또한 물고기와 기타 유기체들은 물의 오염상황에 따라 반응하므로 일부 독성물질이 일시적으로 배출되어 정규적인 수질측정에서는 검출되지 않았다 하더라도, 수중생물체 평가를 통해 이들 오염물질의 영향 유무를 파악할 수 있다. 이와 같이 생물체의 생태적 상황을 분석함으로써 통상의 물리-화학적 측정에 의한 평가보다는 다양한 환경변화를 평가할 수 있으며 이는 수생생물이 수많은 물리적, 화학적 및 생물학적 수질항목들에 총체적으로 반응하기 때문이다.

현재 국내외에 걸쳐서 생물학적 수질 평가를 위해 어류를 포함한 수서곤충(insects), 대형무척추동물(macroinvertebrates) 및 조류(algae) 등과 같은 다양한 생물종들이 지표로 이용되고 있다. 수질생태계의 먹이사슬 단계에서 가장 높은 단계에 위치한 어류는 약 20여종이 독성물질의 누적상황 또는 만성오염의 정도를 대변해 준다. 수질을 평가하기 위해서는 종의 수와 어류의 나이 및 개체군의 구조 등이 분석된다. 수서곤충 및 대형무척추동물을 이용하는 경우 각 생물종의 군집구조를 분석함으로써 수질을 평가할 수 있으나 이들 생물종이 오염에 따른 서식분포의 영향에 대한 연구 자료들을 필요로 한다. 규조류는 다소 오염에 대해 저항능력이 있어 규조류의 밀도를 측정함으로써 수질을 대변해 준다. 규조류의 종 조성은 강마다 다르나 *Navicula radiosa* 와 *Gomphonema acuminatum*가 존재하는 경우 좋은 수질을 나타낸다. 그 외에도 해조류의 서식과 무서식을 이용하여 강과 염습지의 건강상태를 평가할 수 있다. 그러나 수생태계에 바탕을 둔 생물학적 평가방법은 장기간의 방대한 자료의 축적이 있어야 가능하므로 대부분의 국가에서는 널리 적용하지는 못하고 있다.

현재까지 수질을 평가하기 위해 개발된 생물학적 지수들 중 가장 활발히 활용되고 있는 것으로는 IBI(index of biotic integrity), BMWP(biological monitoring working party) 및 DAIPo(diatom assemblage index to organic



pollution) 등이 있다. Biological integrity 개념은 1981년 미국 환경청에서 기존 화학적 수질평가법의 단점을 보완하기 위해 어류를 이용한 수자원 평가도구로 도입되었다. IBI는 시스템의 biological integrity를 정량하고 평가하기 위해 개발되었다(Karr, 1981, 1991). IBI는 어류의 종풍부도(species richness)와 구성(composition), 지표종의 종풍부도와 개체수, 풍부도, trophic organization, function, 산란행동, 개체 상태, 지표 분류군, 잡종(hybrids), 외래종(invasive species) 등을 포함하는 12 metric를 비교함으로써 생태계를 평가한다. IBI score는 12 metric scores의 합으로 나타내며 그 범위는 60(좋음) ~ 12(나쁨)로 나타낸다(Simon and Lyons, 1995). 현재 IBI를 통한 국내외로 연구가 활발히 진행되어 오고 있는데 대형무척추동물(macroinvertebrates)과 조류(algae)를 생물지표로 이용하여 IBI를 평가 및 개발연구가 진행되고 있다(Weigel, et al., 2002; Klemm, et al., 2003).

BMWP는 생물지표로서 대형무척추동물을 이용하여 수질을 평가하는 방법으로 서식처의 수환경을 반영하는 종의 출현과 밀도를 이용하여 평가된다. 저서생물은 이동성이 낮으며 다양한 서식처에 적응함으로써 담수생태계의 구조와 기능을 밝히는데 중요한 생물군으로 인식되었다(곽 외, 2003). 또한 저서생물이 환경변화에 민감하고, 분류군에 따라 비교적 뚜렷한 내성범위를 가지고 있어서 담수생태계의 환경을 평가하는 지표로서 매우 유용하다고 보고되었다 (Hynes and Coleman, 1968; Hellawell, 1986; Resh and Rosenberg, 1984). 이 방법은 대형무척추동물 종(family)마다 오염에 대한 내성이 다르다 것을 기반으로 수질을 평가하는 방법으로, 하루살이과 또는 강도래과의 출현은 수계가 청정하다는 것을 나타내고 내성지수(tolerance score)는 10을 부여한다. 가장 낮은 점수는 지렁이과(Oligochaeta)로 1점이 부여된다. BMWP score는 모든 대형무척추동물 종의 내성지수의 합으로 평가되며 가장 높은 점수는 수질이 가장 양호하다는 것을 반영한다.

그 외 대형무척추동물을 이용하는 지수로는 TBI (trent biotic iIndex) 및 BS (biotic score) 등이 있다. TBI의 경우는 주로 서식처의 특성을 반영하는 종의 출현(강도래류, 날도래류, 하루살이류 등)과 출현 종의 밀도에 따라 지수가 계산되어지는데 점수가 높을수록 양호함을 의미한다. BS는 종(플라나리아, 강도래과, 그물강도래과, 녹색강도래과, 메추리강도래과, 하루살이류, 뱀잠자리류 등)의 출현과 풍부도에 따라 점수가 매겨지며 이를 합산하여 산출된다.

DAIpo는 부착조류를 이용하여 수질을 평가하는 방법으로 유기오염물에 대한 부착조류의 군집지수로 평가된다. 부착조류는 생활사가 짧고, 환경변화에 민감하며 많은 분류군을 가지고 있어 수질에 대한 풍부한 정보를 제공하는 중요한 지표종이다. DAIpo는 오염 민감종과 내성종의 정보를 이용하며 지수의 범위는 0(오탁성) ~ 100(청정성)으로 산출된다(Watanabe, et al., 1986). 그 외에 부착조류를 이용한 수질지수는 수체의 부영양 정도를 평가하는 TDI (trophic diatom index)가 있다. 이 지수는 규조류 종의 오염민감도와 지표값을 이용하는 방법으로 DAIpo와 비슷하나 영양염류와 유기물 오염을 반영한다(Kelly and Whitton, 1995).

### 2.1.3. 통합적 수질평가

인간의 활동을 통한 다양한 여러 화학 오염물질이 호수로 계속 유입되면서 호수의 생태계의 건강을 파괴하여 여러 가지 문제들이 발생하게 되었다. 이런 문제는 생태계뿐만 아니라 사회경제적인 발전이나 인간의 건강에도 영향을 미치게 된다(Cairns 1997, Daily 1997). 1980년 중반이후로 환경을 관리하고 측정하는 자료로서 생태계 건강도에 관한 새로운 수질 지표에 관심을 가지게 되었다(Schaeffer et al., 1988; Costonza et al 1992).

국제적으로 호수의 생태계 건강도 관리 및 평가에 대한 연구와 논문들이 계속 나오기 시작했으며 그중 건강도 평가를 적용할 수 있는 대표적인 지시자들과 방법들이 제시되었다(Xu et al., 2001). 호수 생태계 건강에 적용할 수 있는 대표적인 지시자들이 제안되었는데, 그중에 Hannon에 의해 총체적인 생태계 생산물이 제안되었다(gross ecosystem product(GEP); Hannon, 1985). 또 Rapport et al.(1985)에 의해 index of biotic integrity(ABI)가 제안되었다. 또한 Jorgensen(1995)에 의해서 열역학적 개념에서의 엑서지(Exergy)와 비엑서지(structure exergy)가 수질 지표로 도입하게 되었다.

Xu(2005)는 30개의 이탈리아 호수에 대해 생태계 건강도 평가 방법(ecosystem health index methodology; EHIM)을 적용하여 호수 생태계 건강도를 측정하였다. 이는 기존에 개발된 지표들이 가지는 다양한 조건의 생태계에 적용하기 어려운 문제점을 극복하기 위한 방안으로 제시되었다(Xu, et al., 2001). EHIM 지수는 0~100 까지 지수 범위를 정하였고, 지수 0은 하천의 건강도가 최저이며 지수 100은 건강도가 가장 좋은 상태를 나타낸다. 이 EHIM은 심각하게 오염된 이탈리아의 호수에 성공적으로 적용하였는데 이때 쓰인 지시자들로는 phytoplankton(BA)와 zooplankton(BZ)로 BZ/BA의 비율을 통해 평가되었다(Rast and Holland, 1988).

엑서지와 비엑서지의 개념을 적용한 연구가 국내에서도 진행되고 있다. 이 개념은 하천의 한 시스템 안에서 열역학적 평형상태로부터 거리를 측정하고 이에 따른 생태계의 성장능력을 측정한다. 나아가 동식물성 플랑크톤량을 이용한 엑서지비는 두 종간의 생태학적 균형수준을 평가하게 된다. 비엑서지는 단위 생물량 당 지니는 해당생태계의 정보량을 말하며 생태계 내 조직의 수준이나 종 다양성 수준을 평가할 수 있다. Jorgensen 등은 낙동강의 생태계

건강성분석을 하기 위하여 물금지점의 동식물성플랑크톤 자료를 이용하여 엑서지, 엑서지비(exergy ratio), 비엑서지(structure exergy)를 평가하였다 (Jorgensen, 1995). 엑서지, 엑서지비, 비엑서지를 적용한 결과들은 낙동강 하류인 물금지점에서 Costanza(1992)가 제안한 생태계 건강성 지표들을 적절히 설명하였다. 즉 조류 과다번성 현상을 가진 물금지점의 exergy는 조류량 증대에 따라 증가하는 경향을 보여주고 있어 Costanza에 의해 정의된 활력성(vigor)과 안정성(stability)이 높게 나타났다. 그러나 식물성플랑크톤과 동물성플랑크톤간의 엑서지비는 불균형상태를 나타내었는데, 조류농도가 높아짐에 따라 비엑서지 값이 불안정성을 나타내었다. 따라서 조류가 번성함에 따라서 생태계의 종다양성이 낮아지고 종간에 균형이 파괴되고 있었으며 생태계 구조가 발달하지 못함을 반영하였다. 즉 이 세 지표가 생태계 건강지표로 이용 가능할 것으로 판단되었다.

## 2.2. 수질평가 현황

### 2.2.1. 국외 수질평가 현황

급격히 가속되는 수질오염에 대한 대책으로 수환경 평가 및 이에 대한 제반작업으로 수환경 복원작업이 활발히 이루어지고 있다. 기존의 환경 평가자들은 급속도로 악화되는 수중 생태계 오염을 파악하기 위한 일환으로서 화학적 모니터링을 이용해 왔다(U.S EPA 1991). 그러나 수질 악화는 화학적 오염뿐만 아니라 유로 변경 및 유량 변동, 물리적, 하상구조의 변경 등에 의해 크게 손상되기 때문에, 단순 화학적 평가만으로는 수계 모니터링을 이용한 평가는 많은 시간과 경비가 소요되었고 해석의 모호성 때문에 수질 평가에서 종종 무시되어 왔다(U.S. EPA 1985). 이러한 단점을 보완하기 위해 미국의 환경청에서는 수자원 평가를 위한 기술로서 수중 생태계내의 대표적 trophic composition으로 간주되는 어류, 대형 무척추동물(macroinvertebrate) 및 조류(algae/periphyton)를 생물 지표로 이용한 평가법을 도입하고 있다. 또한 지속적 생태계 관리(sustainable ecosystem management)의 기초가 될 수 있는 통합수질 환경지표 개발을 위한 생태정보 연구는 유럽에서 활발히 진행되고 있는데, EURO project의 하나로써 PAEQUAN(predicting aquatic ecosystem quality using artificial neural network)가 수행되고 있다. 현재 유럽연합국들은 어류나 기타 생물과 이화학적 환경요인을 통합적으로 고려한 지표를 이용하여 수질 개선 및 유지 관리 방안을 모색하고 있다.

미국, EU 등은 1990년대 초반부터 이화학적 지표 중심의 수질평가에서 벗어나 수질영향을 생물학적으로 종합평가하는 방향으로 전환하여 이화학적 지표는 보조지표로 사용하는 추세이다. 생물통합지수를 이용한 평가법은 미국 내 지역적 특성 및 생태적 특성을 고려한 IBI 개념이 미국에 도입된 이래로 미국 내 지역적 특성 및 생태적 특성을 고려한 IBI 수정작업이 지속적으로 이루어지고 있으며, IBI는 1991년에 35개 이상의 주 정부에서 채택되어 하천 평가에 실효를 거두고 있다. 미국의 환경청에서는 하천에 서식하는 생물군집 조사 시 정량화된 채집방법과 하천의 하상구조와 형태학적 특징 등을 이용한 서식지 평가 기법 등을 제시하고 있으며, 이를 근간으로 한 생물(어류, 조류 및 대형 수서 무척추동물) 서식환경 평가를 위해 IBI 평가기법을 핵심도구로 채택하고

있다(EPA, 1996). 유럽연합의 경우 수질관리는 주요 교란물질에 대한 물의 유입 및 유출비율 대한 화학적 수질 검사와 무척추 동물의 조성 및 플랑크톤과 같은 수중 생물들의 종조성에 의해 수질평가가 이루어진다. 일본의 환경성도 영국의 생물학적 평가방법을 참조하여 일본의 실정에 맞는 ASPT(average score per taxon) 지수를 개발하여 현재 적용성에 대해 검토하는 단계이다.

## 2.2.2. 국내 수질평가 현황

국내 수질 및 수생태계 환경기준은 수역별, 항목별로 수질환경기준이 설정되어 있으며, 수역별로는 하천, 호소로 구분하고 항목별로는 생활환경기준, 사람의 건강보호기준으로 구분되어 있다(환경부, 1978). 생활환경기준으로는 pH, BOD, COD, SS, DO, 대장균군수, 총질소 및 총인 등 8개 항목이 있으며, 사람의 건강보호기준으로는 Cd, As, CN, Hg, 유기인, Pb, 6가 크롬, PCB 및 음이온 계면활성제 등 17개 항목으로 구분하고 있다(Table 2.1). 생물학적 수질평가 지표로 저서생물과 어류를 선정하여 수생태계를 평가하고 있다(Table 2.2)(환경부, 2006).

그러나 국내 수질평가방식은 아직까지 BOD, COD 등 이화학적 지표를 이용한 평가방법에 의존하고 있다. 이는 일반오염물질 및 새로운 유해물질이 생태계에 어떤 영향을 미치는지, 또한 먹이사슬을 통해 인체에 어떤 영향을 미치는지 가려내지 못하는 한계가 있다. 국내 수질기준에서 사람의 건강보호를 위한 항목은 카드뮴 등 모두 17개로 미국 126개, 유럽연합 29개, 일본 23개에 비교해 크게 모자란다. 생활환경을 위한 항목도 생물화학적 산소요구량 등 이화학적 항목 5개에 불과하며, 미국 등 선진국에서 70년대부터 적용해온 생물학적 평가는 대장균 항목 하나에 불과하다. 저서생물과 어류를 통해 수생태계를 평가하고자 하나 이는 서식 및 무서식을 통한 평가로서 오염물질이 물속 생태계에 끼치는 복합적 영향을 평가하지 못하고 있다.

또한 일반 수질 항목 조사 분석에서 얻어진 자료의 신뢰성이 부족하며, 숫자적 분석만으로 현 상태를 진단하고 있어 수질항목에서의 수질 측정값이 왜 증가, 감소, 혹은 무 변동에 대한 원인 규명이 불투명하며 과거의 월별 혹은 연별 수질자료를 이용한 현재 및 미래 상태의 진단·예측이 어려운 실정이다.

Table 2.1. Environmental regulation criteria for water quality of the streams in South Korea (Ministry of Environment, 2006).

| <i>Grade</i>                     |                               | <i>Standards</i>  |               |                                   |              |                           |                   |
|----------------------------------|-------------------------------|---|---------------|-----------------------------------|--------------|---------------------------|-------------------|
|                                  |                               | pH  | BOD<br>(mg/ℓ) | SS<br>(mg/ℓ)                      | DO<br>(mg/ℓ) | Coliform<br>(cells/100mℓ) |                   |
|                                  |                               |   |               |                                   |              | Total<br>Coliforms        | Fecal<br>Coliform |
| Daily life<br>environment        | Ia                            | 6.5~8.5   | < 1           | < 25                              | > 7.5        | < 50                      | < 10              |
|                                  | Ib                            | 6.5~8.5   | < 2           | < 25                              | > 5.0        | < 500                     | < 100             |
|                                  | II                            | 6.5~8.5   | < 3           | < 25                              | > 5.0        | < 1,000                   | < 200             |
|                                  | III                           | 6.5~8.5   | < 5           | < 25                              | > 5.0        | < 5,000                   | < 1,000           |
|                                  | IV                            | 6.0~8.5   | < 8           | < 100                             | > 2.0        | -                         | -                 |
|                                  | V                             | 6.0~8.5   | < 10          | No<br>floating<br>solid<br>wastes | > 2.0        | -                         | -                 |
|                                  | VI                            | -   | > 10          | -                                 | < 2.0        | -                         | -                 |
| Protection<br>of human<br>health | For the<br>overall<br>streams | <ul style="list-style-type: none"> <li>- Cd: &lt; 0.005</li> <li>- As: &lt; 0.05</li> <li>- CN: Should not be detected (maximum detection limit 0.01)</li> <li>- Hg: Should not be detected (maximum detection limit 0.001)</li> <li>- Organophosphorus: Should not be detected<br/>(maximum detection limit 0.0005)</li> <li>- PCB: Should not be detected (maximum detection limit 0.0005)</li> <li>- Pb: &lt; 0.05</li> <li>- Cr<sup>6+</sup>: &lt; 0.05</li> <li>- Anionic ABS: &lt; 0.5</li> <li>- Carbon tetrachloride: &lt; 0.004</li> <li>- 1,2-Dichloroethane: &lt; 0.03</li> <li>- Tetrachloroethylene: &lt; 0.04</li> <li>- Dichloromethane: &lt; 0.02</li> <li>- Benzene: &lt; 0.01</li> <li>- Chloroform: &lt; 0.08</li> <li>- Diethylhexylphthalate: &lt; 0.008</li> <li>- Antimony: &lt; 0.02</li> </ul> |               |                                   |              |                           |                   |



Table 2.2. Biological characteristics across the levels of water quality in South Korea (Ministry of Environment, 2006).

| Water quality          | Biological indicators   |  | Habitats and biological characteristics  |
|------------------------|---|--|--|
|                        | Macroinvertebrate   | Fish   |  |
| Very good<br>~<br>Good | Gammaridae,<br>Astacidae,<br><i>Drunella aculea</i> ,<br><i>Cincticostella castanea</i> ,<br><i>kamimuria</i> ,<br><i>Rhyacophila Kua</i> ,<br><i>Glossosma KUa</i> ,<br><i>Limnephilidae</i> ,<br><i>Psilotreta kisoensis lwata</i>                              | <i>Oncorhynchus masou ishikawai</i> , <i>Phoxinus kumgangensis</i> ,<br><i>Brachymystax lenok</i> ,<br><i>Phoxinus oxycephalus</i> ,   | - Very clean water, Fast flow rate<br>- Major benthic substrates: rocks and gravels<br>- Very little attached algae  |
| Good<br>~<br>Moderate  | <i>Semisulcospira libertina</i> ,<br><i>Glossiphonia complanata</i> ,<br><i>Rhoenanthus coreanus</i> ,<br><i>Ephemera orientalis</i> ,<br><i>uracanthella rufta Imanishi</i> ,<br><i>Caenis KUa</i> ,<br>Psephenidae,<br><i>Macronema radiatum</i> ,<br>Mclachlan | <i>Coreoleuciscus splendidus</i> , <i>Zacco temmincki</i> ,<br><i>Plecoglossus altivelis</i> ,<br><i>Siniperca scherzeri</i>           | - Clean water, a little fast or moderate flow rate<br>- Major benthic substrates: rocks and gravels<br>- A little attached algae                             |
| Moderate<br>~<br>Worse | <i>Radix auricularia coreana</i> ,<br>Arhynchobdellidae ,<br><i>Hesperocorixa distanti</i> ,<br><i>Orthetrum albistylum</i>   | <i>Zacco platypus</i> ,<br><i>Opsariichthys uncirostris amurensis</i> ,<br><i>Pseudogobio esocinus</i> ,<br><i>Pseudorasbora parva</i> | - A little turbid water, a little slow flow rate<br>- Major benthic substrates: small gravels and sands<br>- Many green attached algae                       |
| Worse<br>~<br>Worst    | <i>Physa (Physella) acuta</i> ,<br>Tubificidae<br>Chironominae,<br><i>Tinearia alternata</i> ,<br><i>Eristalis tenax</i>  | <i>Carassius auratus</i> ,<br><i>Cyprinus carpio</i> ,<br><i>Misgurnus mizolepis</i> ,<br><i>Silurus asotus</i>                        | - Very turbid water, slow flow rate<br>- Major benthic substrates: sands and silts mostly with black color<br>- Many attached algae with brown or gray color |

### 2.2.2.1. 국내의 수질 지표 개발 현황

낙동강 폐놀오염사건이후 국립환경연구원 호소수질연구소에 실시간 물벼룩 및 어류 경보장치가 도입된 이후, 자체적으로 운용에 관한 연구를 해 왔으나, 괄목할 만한 성과는 발표되고 있지 않다(최, 1997). 한국화학연구원에서는 잉어의 호흡 및 움직임을 이용한 독성측정장치를 제작, 유해물질의 독성을 평가하였다(한국화학연구소, 1995). 전극을 이용하여 잉어의 호흡과 움직임을 측정할 수 있는 장치를 구성하여 생물학적 반응을 감지할 수 있는가를 평가하였다. 4 종류의 화학물질을 이용하여 이 장치의 측정 감도, 응용성 등을 평가한 결과 급성적 농도수준에서는 적어도 3시간 이전에 감지가 가능하였고, 카드뮴의 경우는 급성독성치의 30분의 1 수준도 감지할 수 있었다. 따라서 국내에서도 이러한 장치의 개발 가능성을 확인할 수 있었다(한국화학연구원, 1996). 현대건설 기술연구소에서는 ‘어류를 이용한 수질상시감시 및 경보장치’를 개발 현재 화승산업(주)에서 판매하고 있다. 부산대학교 전태수 교수팀은 1998년부터 과학재단의 지원으로, ‘인공신경 회로망 및 수리적 방법을 통한 주요생물종의 독성물질에 대한 행동반응 유형화’에 대한 연구를 수행하면서, 어류 및 수서무척추동물인 chironomidae에 대한 독성물질 반응에 대한 기본적인 연구를 수행한 바 있다.

상시감시체제는 아니지만, 미생물을 이용한 독성물질 감지에 대한 연구는 상대적으로 활발하여, 광주과학기술원의 구만복 교수팀(1999, 2000)은 유전공학적으로 조작한 생물발광박테리아를 이용한 minibioreactor system을 개발하였다. 서울대학교 김상중 교수팀(1999)은 환경부 G-7과제의 연구성과로, 발광미생물을 이용한 유해물질 검출장치를 개발을 하였으나, 실용화에는 미치지 못한 실정이다.

생물의 생태 해석을 위해 인공신경지능망인 Kohonen network를 이용하여 생태군집을 분류하고(Chon 등, 1996), 나아가 군집의 시간적 변이에 대한 유형화 연구가 시도되었다(Chon 등, 2000). 하지만 국내외의 연구는 어류 및 대형 생물에 국한되어있다(이 등, 2004; 김과 안, 2004). 어류보다 민감도 및 특이성에서 우수한 미생물을 이용한 통합적 수질지표개발연구는 초기 단계이며(장과 이, 2003; 최와 신, 1997; 김 등, 2002), 미생물을 이용한 통합수질 지표개발은 거의 없다.

우리나라에서는 1960년대 후반부터 저서성 대형무척추동물의 군집분석, 현존량 조사 및 수질관정이 많이 실시된 바 있으며 산업발전으로 인해 오염된 하천에서의 저서성 대형무척추동물 군집구성과 수질평가에 관한 연구가 다수 있었다(서와 박, 1982; 나 등, 1986). 특히 90년대 후반부터 군집분포 및 동태의 생태적 관점에서 많은 연구가 진행되었다(배 등, 1996; 윤 등, 1998). 그리고 수질등급에 따라 중요 출현 분류군을 조사한 바 있다(Yoon, 1995). 또한 90년대 중반 이후 건설기술연구원등에서 친생태적 하천회복공법 등이 연구되었다(박종화, 1997). '90년대에 생태자료에 대한 인공신경망 적용 연구를 수행하였는데 부분적으로 국제적 경쟁성을 확보하고 있는 분야가 있다. 특히 국내에서 1996년 SOM을 생태자료에 최초로 적용하여 data mining을 수행하였는데(Chon et al., 1996), 이는 2000년 이후에 국제적으로 생태분야에 적용되고 있다. 또 Recurrent network도 2001년에 최초로 적용하여 시계열 자료에 대한 해석이나 예측도 경쟁성을 확보하고 있다(Chon et al., 2001). 기타 Kohonen network이나 Adaptive resonance network을 이용한 hybrid network을 적용한 연구도 수행된 바 있다(Chon et al., 2000). 그러나 이러한 생태적 자료에 대한 계량화 작업에도 불구하고, 국내의 경우 외국의 경우와 같은 통합적 수질지표개발과 에너지흐름을 유기적으로 조사한 경우는 드물다.

#### 2.2.2.2. 국내 수질환경관리를 위한 통합적 수질지표개발의 필요성

국내에서는 오염하천의 관리를 위한 통합적 수질지표개발과 에너지흐름을 유기적으로 조사한 경우는 드물다. 주요 분류군 및 주요 생태계 등에 대해서 독립적으로 연구가 진행되었다. 이를테면 생물평가방법에 의거하여 국내에서도 수질등급에 따라 중요 출현 분류군을 조사한 바 있다(수서곤충검색도설; 1995). 그러나 이는 종의 유무에 의한 수질판정이므로 분해능과 에너지 흐름을 다루지는 않는다.

국내의 생물평가 연구는 주로 군집연구에 국한되어 있다. 그러나 통합수질지표 개발은 군집의 구조적 측면뿐만 아니라 기능적 측면에서도 시도되어야 한다. 각종생물의 구조유전체학 및 기능유전체학에 관련한 여러 기술이 급격하게 개발되고, 상품화되고 있다. 이들 기법을 이용하면 생태계의 기능과 지속성에 있어서 미생물의 역할규명이 크게 가능해질 전망이다. 특히, 미생물 군집구조와 구성성분, 독성물질의 유입과 농업과 기타산업 활동 등에 의한 환경변화에 대한 그들의 반응과 적응을 효율적 분석할 경우, 환경에 생태 및 환경문제의 궁극적 해결책을 제시할 수 있을 것이다.

또한, 통합 수질 지표 개발 연구의 결과로 일반적인 생물군집의 구조적 측면에서 지표를 제시할 수 있다. 현재까지 생물지표개발에 있어서 분해능에 대한 연구는 드물었다. 그러므로 분해능과 에너지 흐름에 대한 연구는 수환경에 관여하는 생물-비생물의 제반 중요요인을 내재시키어 종합적으로 나타낼 수 있으므로 다면적이고도 입체적으로 하천 생태계를 표출할 수 있다.

오염하천 생태계의 복합성에 관한 연구는 우리나라 뿐 만이 아니라 국제적으로도 시작하는 단계에 있다. 앞으로 이 분야가 발전하여 통합 수질지표가 제시되면 오염하천 생물 콤플렉스의 구조와 기능을 생태학적으로 검토할 수 있을 것으로 판단된다.

### 2.2.2.3. 통합 수질지표 개발 방향

수환경을 개선하는 생물의 능동적 반작용은 일찍부터 자연정화의 하나로 인식되어 왔으나, 그 기능을 인위적으로 극대화하려는 시도는 1950년대 말 이후에 구체화되었다. 이러한 생태공학적 기술은 넓은 의미로는 하·폐수처리를 통한 오염원관리(source control) 기술과 자연수면에서의 생물관리 기술로 대별되고 있다. 자연상태에서 생물군집의 조절에 의한 수질관리에 관한 연구는 1950년대 말을 전후하여 주로 유럽을 중심으로 시작되었는데, Shapiro(1975)는 이를 biomanipulation이라 명명하고 “biomanipulation이란 수질개선을 위한 자연산 수생생물 군집의 관리”라고 정의한 바 있다. biomanipulation은 환경오염이 인간에 의하여 생태계가 환경에 이상적(異常的)으로 반응하는 생태계 대사(metabolism)의 병리적인 문제로서 해석하고 생태계 ‘체제’의 관점에서 해결책을 유도하는 것이다. 수질관리를 위한 지속적 생태계 관리와 환경보전은 생태적인 관점에서 방향설정을 해야 한다. 그러나 실제 환경관리 현장에서는 환경관련법에 나타나 있는 기준에 의거하여 일정한 기준을 넘지 않도록 규제하는 방향으로 시행되고 있고, 인간의 위생적인 측면만 강조될 뿐 생태계 대사나 전반적인 추이는 고려하지 않고 있다. 관리기준도 pH, DO, BOD, 부유물질과 중금속의 함유도 등의 대부분 이화학적 요인에 의거하여 결정된다.

생태계(예, 하천, 삼림)의 보전 대상은 생태계의 주요 구성요인인 생물 군집이다. 생태계 관리의 주된 과제는 생태계 건강성(ecosystem health)을 위해서 인간을 포함한 생물군집의 양태를 전체적으로 고찰하는데 있다. 관리지표에 생물 및 생태적 방법을 포함하고 이화학적 지표와 상보적으로 적용하여야 효과적인 수질의 진단이 가능하다. 생태적 수질평가는 수환경에 관여하는 생물-비생물의 제반 중요요인을 내재시키어 종합적으로 나타낼 수 있으므로 다면적이고도 입체적으로 하천 생태계를 표출할 수 있다.

그러나 생태계의 제반 자료는 복잡하고 비선형적 이어서 전통적인 통계법으로는 일반적으로 분석하기 힘든 점이 많았다. 생태정보학의 주요 방법론인 인공신경망은 융통성과 자체조직능력이 있어 적응과 진화적인 현상을 본질적으로 내재하고 있으므로 복잡한 생태현상에 적용해 볼 수 있다. 앞으로 수환경이 행정적이거나 법률적인 규제의 차원으로부터 적극적이고도 능동적으로

생태계를 지속적으로 관리하거나 복원하고, 생태계 건강성을 적극적으로 추구하는 차원으로 옮겨가는 경향성을 감안할 때, 생태적 수질평가는 매우 의미가 있다 하겠다. 생태정보학을 통한 생태적 수질평가는 제반 생태공학적 정보를 줄 수 있고, 국제적 생태계 관리전략 연구에 핵심적으로 이용될 수 있다.

그러나 현재까지 대부분의 연구는 조류(생산자), 대형무척추동물(소비자) 및 미생물(분해자) 군집별로 독립적으로 연구가 진행되어온바, 생태적 수질평가 및 관리를 위해서는 복잡한 생태환경조건에서 이루어지는 이들 군집의 구조와 기능을 통합하는 생태적 지표 및 지수개발이 필요하다.

## 제 3장 미소 생태계 환경 및 생물요인

본 연구에서는 하천 생태계 관리를 위한 이화학·생물학적 통합수질지표 개발을 위한 기초자료를 수집하기 위해 미소생태계의 환경요인과 생물요인을 분석하였다. 이화학적 환경요인은 기온, 수온, pH, conductivity, DO, BOD, 탁도, 수심, 유속, COD, TOC T-P, T-N 및  $\text{NO}_3^-$ -N 등 기본적인 이화학적 항목들과 하천에 서식하는 조류(생산자), 저서성 대형무척추동물(소비자)과 biofilm을 형성하는 미생물(분해자)의 군집구조를 조사하였다. 오염의 종류 및 정도에 따른 미소생태계의 변화를 측정하기 위해 오염원의 종류를 다양하게 선정하였다.

### 3.1. 실험방법

#### 3.1.1. 조사지점

산업화 사회에서 오염의 증가와 오염원의 증가로 인하여 생활 주변에 산재해 있는 대표적 오염의 종류가 증가되었다. 오염원의 종류는 사회적으로 중요시되고 있는 가정하수(유기물), 산업폐수, 축산오수 및 농업(경작지)오수를 선정하였다. 오염의 종류 및 정도를 비교할 수 있는 대조군으로 중요 오염원이 유입되는 하천의 상류(일부 지점은 제외)를 선택하였다.

조사지점은 부산 및 경남지역으로 총 5개 지역(오염원에 따른 분류), 21곳의 조사지점을 선정하였다(Fig. 3.1, Table 3.1). 조사기간은 2002년 1월부터 2004년 1까지 Duk, Ddk, Dkh Dks, Dag 등 5개 지점에 대해 월별(19개월), 2003년 3월부터 2004년 12월까지 ReK, ReA, ReT, OCU, ONS, DUK, FaC, FaA, LiT, LiK, InJ, HJD 등 12개 지점에 대해 계절별(7계절), 2005년 2월부터 2007년 4월까지 DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD 등 8개 지점은 계절별(10계절)로 조사를 수행하였다. 측정 항목별 자료행렬 수는 환경요인의 경우 13 수질항목 × 253 지점수, 대형무척추동물 85 family × 253 지점수, 조류는 60 species × 253 지점수, 미생물 115 species × 253 지점수의 형태를 가진다.

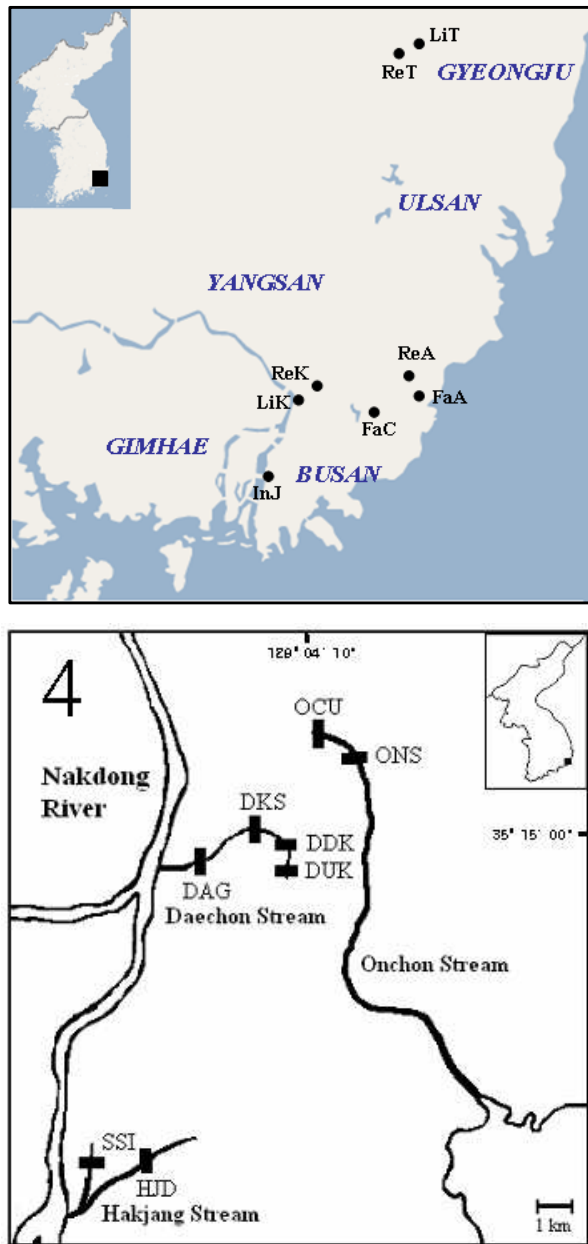


Fig. 3.1. Map of the sampling sites for this study.  
(refer to Table 3.1 for the details)



Table 3.1. Sampling sites for this study.

| <i>Sampling period</i>                | <i>Pollution source</i>   | <i>Stream sites</i>       |
|---------------------------------------|---------------------------|---------------------------|
| Jan. 2002 ~ Jan. 2004<br>(19 months)  | Reference<br>(pristine)   | Duk (Daechon stream up)   |
|                                       | Domestic sewage           | Ddk (Daechon Stream mid)  |
|                                       |                           | Dks (Daechon Stream down) |
|                                       |                           | Dkh (Daechon Stream down) |
| Restoration stream                    | Dag (Daechon Stream down) |                           |
| Mar. 2003 ~ Dec. 2004<br>(7 seasons)  | Reference<br>(pristine)   | ReA (Ilkwang up)          |
|                                       |                           | ReT (Toham up)            |
|                                       |                           | ReK (Kumkok Stream up)    |
|                                       |                           | OCU (Onchon Stream up)    |
|                                       | Domestic sewage           | DDK (Daechon Stream mid)  |
|                                       |                           | ONS (Onchon Stream mid)   |
|                                       | Farming                   | FaC (Chulma mid)          |
| FaA (Ilkwang mid)                     |                           |                           |
| Livestock                             | LiT (Toham up)            |                           |
|                                       | LiK (Kumkok Stream mid)   |                           |
| Industry                              | HJD (Hakjang Stream down) |                           |
|                                       | InJ (Jooksung down)       |                           |
| Feb. 2005 ~ Apr. 2007<br>(10 seasons) | Reference<br>(pristine)   | DUK (Daechon Stream up)   |
|                                       |                           | OCU (Onchon Stream up)    |
|                                       | Domestic sewage           | DDK (Daechon Stream mid)  |
|                                       |                           | DKS (Daechon Stream down) |
|                                       |                           | ONS (Onchon Stream mid)   |
| Restoration stream                    | DAG (Daechon Stream down) |                           |
| Industry                              | HJD (Hakjang Stream down) |                           |
|                                       | SSI (Sasang Stream down)  |                           |

### 3.1.2. 미소생태계의 이화학적 환경요인

생물조사와 병행하여 각 조사장소에서 기온, 수온, pH, conductivity, DO, BOD, 탁도, 수심, 유속, COD, TOC T-P, T-N 및  $\text{NO}_3^-$ -N 등 기본적인 이화학적 항목들에 대해 조사를 수행하였다. 기온, 수온, pH, 전도도(conductivity) 및 DO는 수표면 하 20cm 되는 수층(water column)를 대상으로 전기적 측정법(Harris Environmental Test Center, Philip Harris Ltd.)으로 조사하였다. 탁도는 spectro-photometric methods(LaMotte Inc, 2020)를 사용하였고, 유속계(Model 2100-STDx)를 사용하여 유속을 측정하였다. 하상구성물은 Cummins와 Lauff (1969)에 따라 10cm 이상에서부터 0.0625mm까지 12단계의 크기로 나누어 조사하였다. BOD, 탁도, 수심, 유속은 3회 반복 측정하였다.

조사지점의 시료로부터 유기물질의 농도를 측정하기 위하여 COD, BOD, TOC, DOC 등을 측정하였다. BOD의 경우 시료를 5일간 배양 후 Standard Methods(APHA, 1985)에 따라 분석하였고, COD의 측정은 Standard Methods(APHA, 1985)에 의하여 Closed Reflux, titrimetric method로 분석하였다. TOC와 DOC량은 연소/비분산 적외선 가스 분석법에 기초를 둔 TOC-5000을 이용하여 측정하였고 DOC는 시료를 GFC filter로 여과한 후 측정하였다.

무기영양염류인 Total phosphorus(T-P), Nitrate ( $\text{NO}_3^-$ )는 각 조사지점에서 하천수를 채취하여 standard colorimetric techniques에 따라 분석하였고(APHA, 1985), total nitrogen(T-N)은 킬달 증류장치를 이용하여 매뉴얼에 따라 수행하였다.

### 3.1.3. 미소 생태계의 군집구조

#### 3.1.3.1. 저서성 대형무척추동물 군집구조

저서성 대형무척추동물의 채집 및 분류는 부산대학교 생명과학부 생태 및 행동시스템실험실(전태수 교수)에서 수행되었다. 각 조사지점 내 미소생태계의 군집조사를 위해 시료는 Corer 채집기(직경 10 cm, 길이 15 cm)를 이용하여 3 ~ 5회 반복 정량 채집하였다.

채집된 저서성 대형무척추동물은 formalin 7%에 고정시켜 실험실에서 현미경(Olympus SZH10) 하에 동정(분류)하였다. 일반적인 수서곤충 및 저서생물의 분류는 윤(1995)과 Brigham 등(1982), Merritt와 Cummins(1996), Pennak(1978), Brinkhurst(1986) 등을 따랐다.

#### 3.1.3.2. 부착조류 및 부유조류 군집구조

부착조류 및 부유조류의 채집 및 분류는 전남대학교에서 수행되었다. 부착조류의 채집은 하상의 편평한 돌의 상부를 5 cm × 5 cm를 구획하여 부드러운 솔로 수세한 후, 현장의 물을 이용하여 폴리에틸렌 용기(100 ml)에 넣어 Lugol's iodine 용액(5-7%)으로 고정하여 실험실로 운반하였다. 채집, 고정된 시료의 동정, 분류는 농축된 시료로부터 표본을 제작하여 광학현미경(Nikon, 1000배 이하) 하에서 실시하였으며, 분류체계는 Hirose *et al.* (1977)과 Hustedt (1930)에 따랐다.

시료를 적당량 취해 Sedwick-Rafter(S-R) chamber 에 올려놓고 커버 글라스로 덮은 후, 광학현미경(Axioplan, Zeiss, Germany)을 이용하여 고배율(X400, X1000) 하에서 계수, 동정하여 ml 단위로 환산하였다.

### 3.1.3.3. 미생물 군집구조

미생물의 채집은 하상의 퇴적물을 10cm core를 이용하여 채취하였으며, dry ice를 이용하여 냉동 보관하여 실험실로 운반하였다. 채집된 시료의 동정은 총 DNA 추출 후 PCR-DGGE 기법으로 16S rDNA의 염기서열을 분석함으로써 실시하였다.

본 연구에서는 하천의 오염 정도에 따라, 미생물군집의 구조를 분석하기 위하여 각 조사지점의 하천 저니 샘플로부터 DNA를 추출하여 PCR-DGGE 기법을 이용하였다. PCR-DGGE 기법을 이용한 오염하천 미생물의 군집구조 분석은 Fig. 3.2에 제시된 바와 같이 수행하였다.

각 지점 별 시료에서 total DNA를 추출은 FastDNA Spin Kit(Bio101, USA)를 이용하였으며, 16S rDNA 증폭을 위한 주형으로 이용하였다. 증폭된 16S rDNA 증폭 산물(100 ng)은 16S rDNA V3 region 증폭의 주형으로 재이용되었으며, 16S rDNA V3 region 증폭에 사용된 primer는 40bp의 GC-Clamp가 포함된 GC-341F(5'-CGC CCG CCG CGC GCG GCG GGC GGG GCG GGG GCA CGG GGG GCC TAC GGG AGG CAG CAG-3')와 518R(5'-ATT ACC GCG GCT GCT GCT GG-3')을 Bioneer Inc.(Daejeon, Korea)에서 주문·제작하여 사용하였다. 16S rDNA V3 region은 touchdown PCR을 수행하여 증폭하였으며, 반응조건은 95°C에서 5분간 초기 열처리 후, 95°C에서 30°C 동안 변성시키고, annealing 온도는 초기에 65°C에서 시작하여 매 2cycle마다 1°C씩 감소되도록 설정하고, 55°C가 되면 15cycle을 더 수행하도록 하여 30초간 반응시키고 신장을 위하여 72°C에서 45초 반응 후 최종으로 72°C에서 10분간 처리하고 반응을 마쳤다. PCR 산물은 1.0%의 아가로스 겔에서 전기영동을 수행하여 확인하였다.

Touchdown PCR을 수행하여 얻은 PCR 산물은 BioRad Dcode System (BioRad Laboratories, USA)으로 DGGE를 수행하였다. Denaturing gradient gel은 농도구배(40%~70%)가 연속적으로 형성되도록 하였다. Touchdown PCR을 통해 얻은 산물 20  $\mu$ l를 loading하여 1X TAE 완충용액에서 60°C, 60V로 16 ~ 18시간 전기영동 하였다. 전기영동이 끝난 gel은 EtBr을 사용하여 염색 후 Scientific Image System(Kodak, USA)로 관찰·분석하였다.

Denaturing gradient gel상에서 위치가 다른 밴드로부터 DNA 단편을 회수하

기 위하여 먼저 각각의 밴드를 자른 뒤 3차 증류수로 세척 후, TE 완충용액 25 $\mu$ l를 첨가하여, 원심분리(14,000 $\times$ 1min)하고 -70 $^{\circ}$ C에서 5분, 50 $^{\circ}$ C에서 5분을 세 번 반복한 후, 마지막 원심분리(14,000 $\times$ 1min)를 하여 상등액을 취하였다. 각 밴드에서 회수한 DNA에 대하여 341F와 518R을 가지고 PCR을 수행하여 재증폭 하였다. 재증폭 산물은 정제하여 cloning(Promega, USA) 후 염기서열을 분석(Cosmo Genetech, Korea)하였다. 이로부터 얻어진 염기서열 데이터들을 NCBI BLAST search를 통해 가장 유사도가 높은 미생물 종을 밝혀내고 발표된 문헌을 통하여 그 특성들을 파악하였다.

또한 PCR-DGGE 기법을 이용하여 분석한 결과 나타난 band를 디지털 이미지 분석 장치를 이용하여 디지털 값으로 환산한 후 미생물의 밀도로 간주하였으며, 이 값을 SOM 분석에 활용하였다.

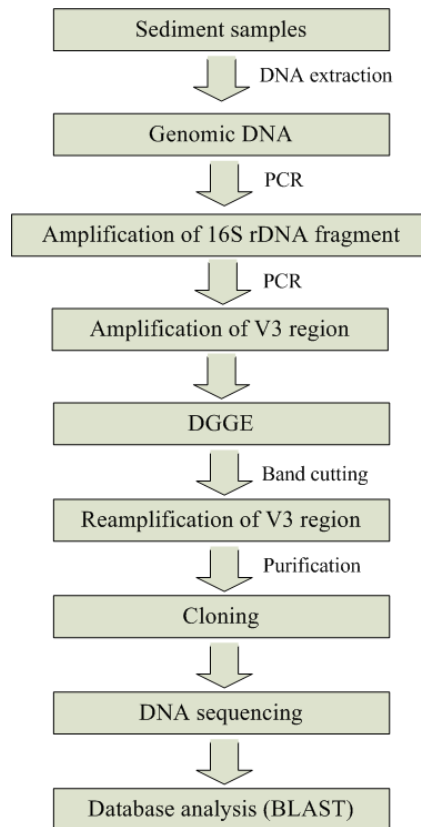


Fig. 3.2. Procedure for analysis of microbial communities using PCR-DGGE.

## 3.2. 결과 및 고찰

### 3.2.1 미소생태계의 이화학적 환경요인

#### 3.2.1.1. 수온, 유속 및 수심

각 분류군의 서식처 특성을 반영해주는 물리적 환경요인인 수온, 수심, 유속을 측정하였다.

수온은 연평균 10 ~ 17℃로 조사되었으며(Fig. 3.3), 계절별 평균기온은 봄 조사 시 3 ~ 5 ℃, 겨울 조사 시에는 3 ~ 7.5 ℃, 여름조사 시 18 ~ 28 ℃의 분포를 보여 각 계절의 특성을 잘 반영하였다. 각 조사지점 간의 온도차는 조사시기가 오전, 오후로 다양하게 이루어졌으며, 시료 채취지점의 위치에 의해 일조량이 다르기 때문이다.

평균유속의 경우 연 평균 0 ~ 49.1 cm/s로 조사되었다(Fig. 3.4). 전체적으로 겨울 조사가 봄에 비해 평균유속은 느리게 나타났다. ReK와 같이 간헐적으로 물이 흘러 고인 곳이 많은 경우 유속이 낮게 나타나며, 잦은 강우로 인해 갑작스런 유입이 발생하는 경우 유속이 높아지는 것으로 나타났다.

수심은 연 평균 6.67 ~ 24.4 cm로 나타났다(Fig. 3.5). 조사지점 중 산업폐수 오염지역인 SSI 지점의 수심이 가장 낮은 것으로 나타났다. 또한 수심은 강우 및 지형적인 특성에 의해 영향을 많이 받는다.

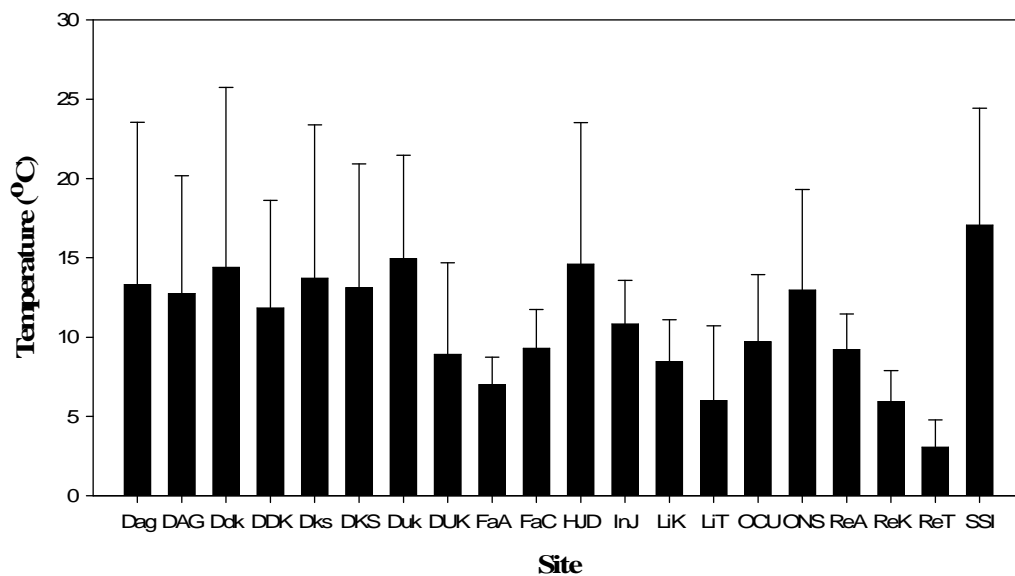


Fig. 3.3. Average temperature of water samples from the sampling sites over the entire sampling period.

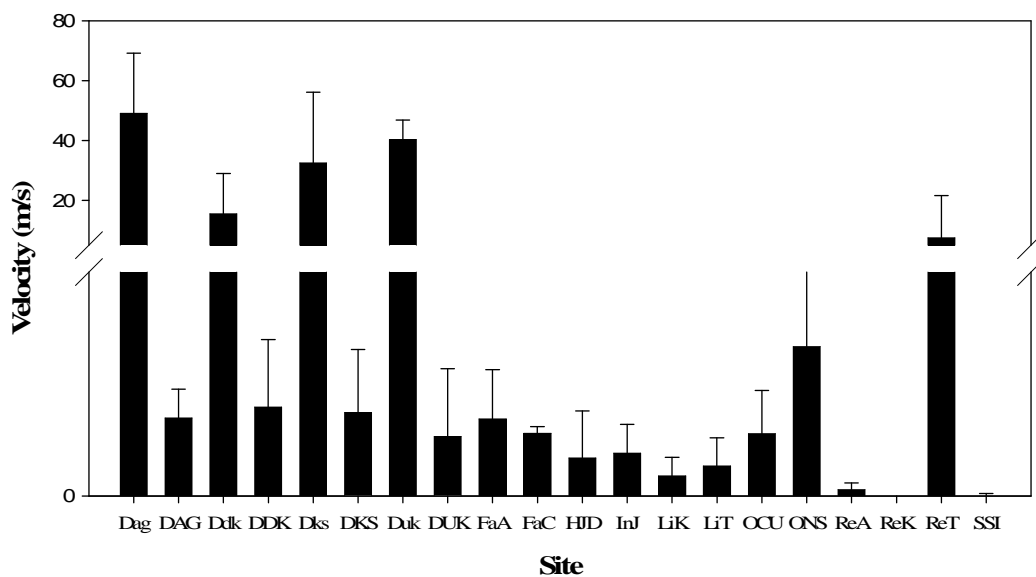


Fig. 3.4. Average velocity of stream water from the sampling sites over the entire sampling period.

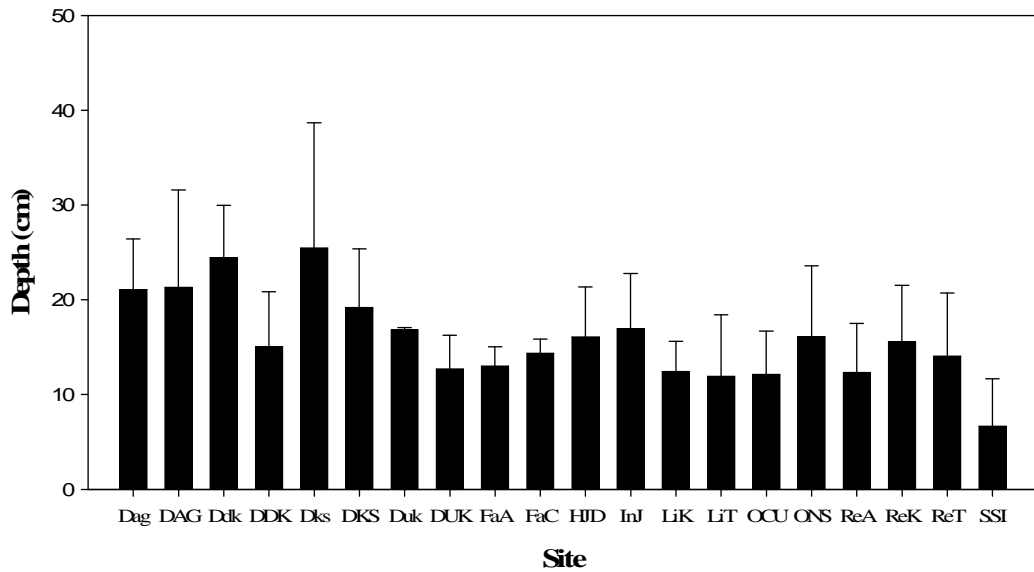


Fig. 3.5. Average depth of stream water from the sampling sites over the entire sampling period.



### 3.2.1.2. pH

조사지점에 대한 pH 연평균 측정치는 5.45 ~ 7.82로 나타났다(Fig. 3.6). 전체적으로 중성 pH를 띠는 것으로 나타났으나 산업폐수 오염지역인 SSI 지점에서 pH 5.45로 가장 낮은 수치를 보였다. 이는 조사지 주변 산업단지로부터 유입되는 폐수의 성상에 의한 것으로 사료된다. 대조조사지의 경우 타 지점들에 비해 비교적 낮은 pH를 나타내었으며, 특히 Duk 지점과 ReT 지점에서의 pH가 각각 6.56과 6.79로 나타났다.

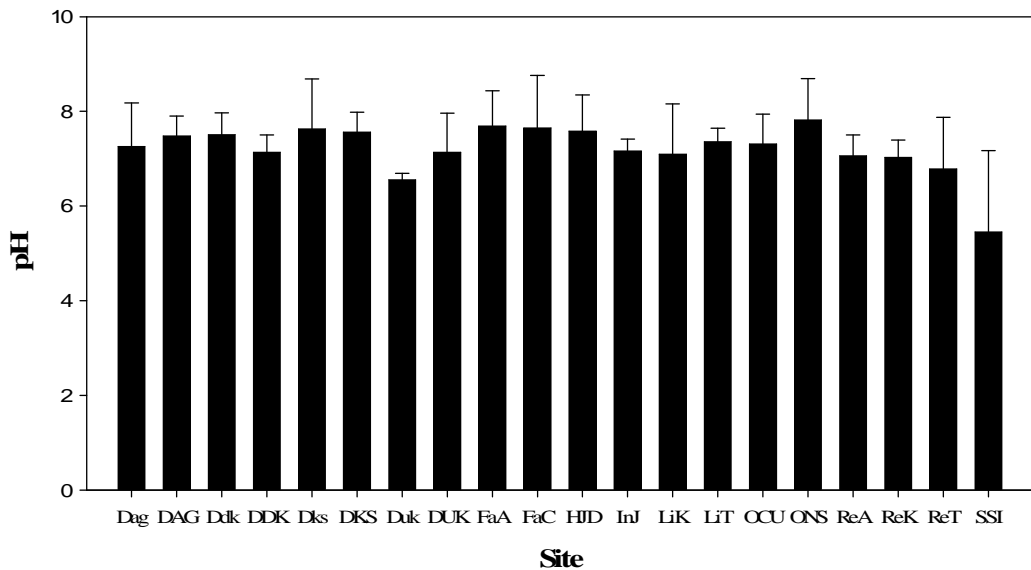


Fig. 3.6. Average pH of water samples from the sampling sites over the entire sampling period.

### 3.2.1.3. Conductivity

전기전도도는 전체적으로 연평균 35.7 ~ 2259.6  $\mu\text{S}$ 의 범위로 넓게 조사되어 다양한 오염범위를 나타내었다(Fig. 3.7). 대조 조사지(청정지역)의 경우 전체조사기간을 통해 35.7 ~ 183.9  $\mu\text{S}$ 로 나타나 ReK(183.9  $\mu\text{S}$ ) 지점을 제외하고는 다른 오염지역에 비해 낮은 측정치를 보였다. 반면, 유기물오염지역은 전체적으로 137.4 ~ 342.5 $\mu\text{S}$ 으로 나타났으며 ONS 지점의 측정치가 전반적으로 높게 나타났다. 축산폐수 및 농경지 오염지의 경우 유기물 오염지역의 수치와 유사한 측정치를 보였다. 산업폐수 오염지인 SSI와 InJ의 경우에는 각각 2259.6, 1374 $\mu\text{S}$ 으로 전체 조사지점에서 가장 높은 수치를 나타내었으며, 조사기간을 통해 전반적으로 가장 높게 나타났다.

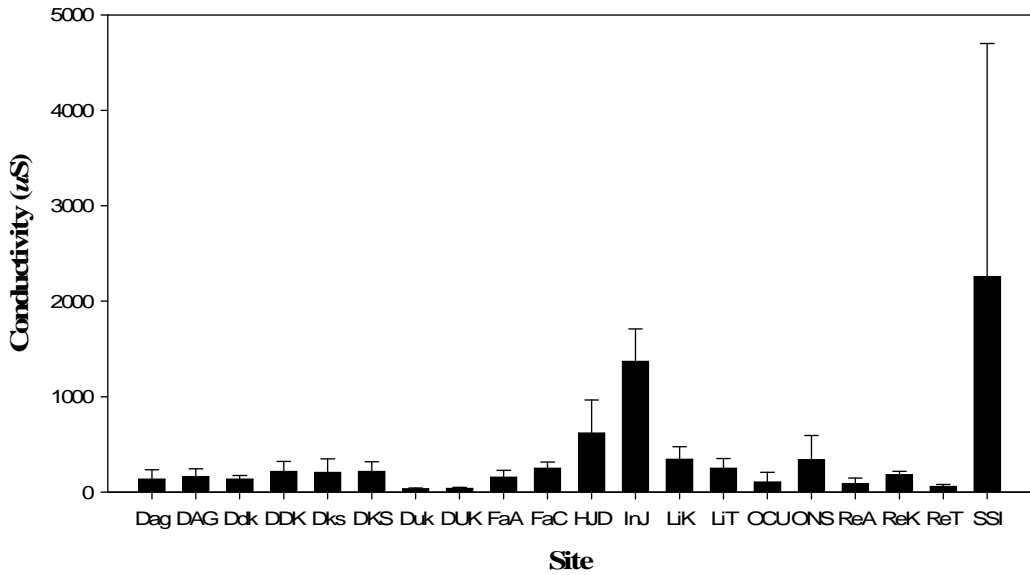


Fig. 3.7. Average conductivity of water samples from the sampling sites over the entire sampling period.

### 3.2.1.4. Turbidity

탁도는 전체적으로 연평균 1.03 ~ 216.8 NTU의 범위를 보여 다양한 오염구배를 나타내었다(Fig. 3.8). 대조 조사지의 경우 1.47 ~ 6.74 NTU로 낮았으며, 큰 변이를 보이지 않았다. 유기물 오염지역인 Ddk, Dks, Dag 지점은 44.78, 42.23, 27.26 NTU로 비교적 높게 나타났다. 이는 시료채취 기간 중 하수관로 공사로 인해 탁도가 높아졌기 때문인 것으로 판단된다. 회복지인 DAG의 경우 대조조사지보다 낮은 1.03 NTU의 측정치를 보였다. 농경지 오염지역인 FaC는 연 평균 탁도가 76.73 NTU로 높게 나타났으며, 특히 2003년 봄에 217.67 NTU로 매우 높게 나타났는데, 조사기간에 인근 농경지로부터 유입된 탁류에 기인하였다. 산업폐수 오염지역에서는 점차로 탁도가 증가하였다. 특히 SSI의 경우 연 평균 측정치가 216.84 NTU로 전체 조사지점 중 가장 높은 측정치를 나타내었다. 이 지점에서의 높은 측정치는 산업폐수뿐만 아니라 생활오수의 영향도 큰 것으로 사료된다.

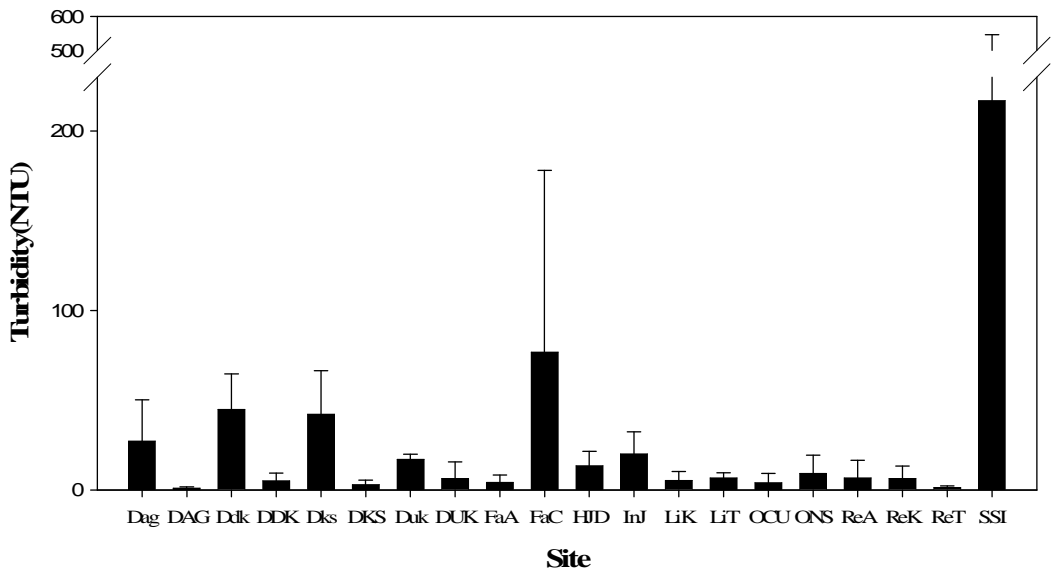


Fig. 3.8. Average turbidity of water samples from the sampling sites over the entire sampling period.

### 3.2.1.5. Biochemical oxygen demand

전체적으로 BOD<sub>5</sub>도 넓은 범위에서 나타났다(Fig. 3.9). 전 지점에 걸친 연평균 BOD<sub>5</sub> 측정치는 0.8 ~ 76 mg/L 조사되었다. 대조 조사지의 경우 0.8 ~ 4.74 mg/L로 낮은 오염농도를 보였다. 유기물 오염지역은 4.66 ~ 46.51 mg/L로 다양한 오염구배를 나타내었으며 대천천 하류 지점인 Ddk 지점이 가장 높은 오염도를 나타내었다. 회복지인 DAG의 경우 2.54 mg/L로 대조조사지 정도의 오염도를 나타내었다. 산업폐수 오염지역의 경우 SSI 지점과 InJ 지점이 각각 49.46 mg/L와 76mg/L으로 높은 오염도를 나타내었으며 InJ 지점이 가장 심하게 오염된 것으로 나타났다.

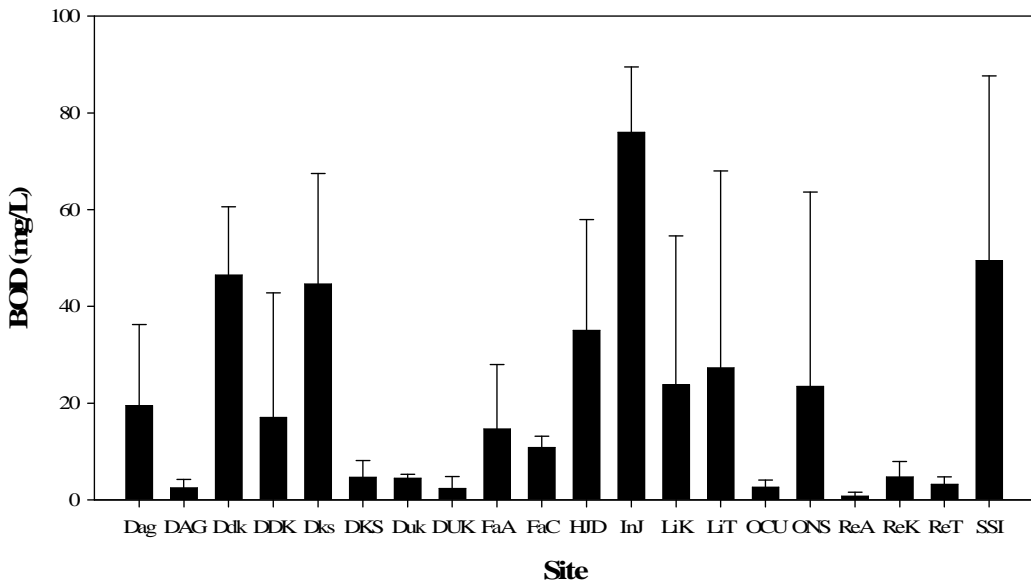


Fig. 3.9. Average biochemical oxygen demand of water samples from the sampling sites over the entire sampling period.

### 3.2.1.6. Total organic carbon

하천에 유입되는 유기물량의 조사를 위한 총유기탄소량(TOC) 측정결과 연평균 0.80 ~ 516.89 mg/L로 나타났다(Fig. 3.10). 대조조사지의 경우 타 오염지역에 비해 TOC 측정치가 낮게 나타났다. 농경지오염지역, 축산폐수 오염지역, 그리고 산업폐수 오염지역이 다른 오염원에 비해 TOC가 높게 나타났는데, 특히 산업폐수 오염지역의 InJ 지점과 SSI 지점의 측정치가 각각 30.24 mg/L와 516.89 mg/L로 SSI 지점의 오염도가 가장 높았다. 그리고 회복지인 DAG는 0.94 mg/L로 대조조사지 측정치와 유사한 수준의 TOC 오염도를 보였다.

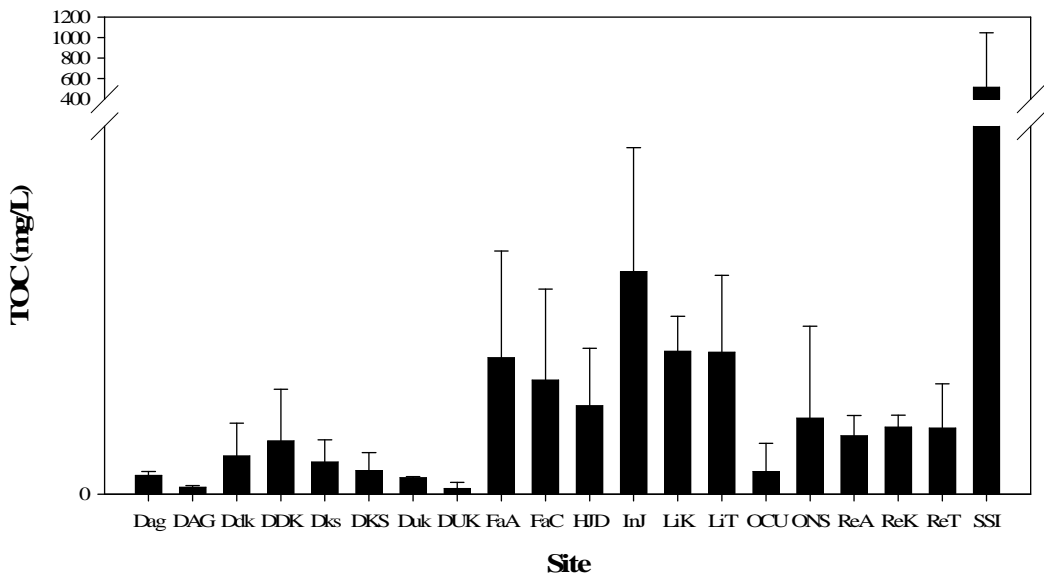


Fig. 3.10. Average total organic carbon of water samples from the sampling sites over the entire sampling period.

### 3.2.1.7. Total phosphorus

오염원의 종류에 따른 특성을 조사하기 위해 총인(T-P)의 양을 측정하였다. 전체적으로 총인의 연평균 측정치는 0.01 ~ 25.11 mg/L로 다양한 분포를 보였다(Fig. 3.11). 산업폐수 오염지역인 HJD지점과 SSI지점이 각 5.02 mg/L, 25.11mg/L로 가장 높게 나타났으며 오염지역별로 살펴보면 산업폐수 오염지역인 SSI가 가장 높았다. 유기물 오염지역의 경우 Ddk, Dks 지점이 3.37 mg/L, 4.85 mg/L로 산업폐수 오염지역인 SSI 지점을 제외하고 유기물 오염지역 중에서도 비교적 높은 오염도를 보였다. 특히 유기물 오염지역의 오염변화 추이를 살펴보면 여름철에 총인 오염도가 급격하게 상승하는 것을 알 수 있었다. 이는 계절적 특성에 의한 것으로 인구밀집 지역의 가정하수 내에 포함되어 있는 합성세제 및 음식물에 의한 영향으로 사료된다. 회복지인 DAG의 경우 타 수질항목들과 마찬가지로 0.46 mg/L로 대조조사지 정도의 오염도를 보였다.

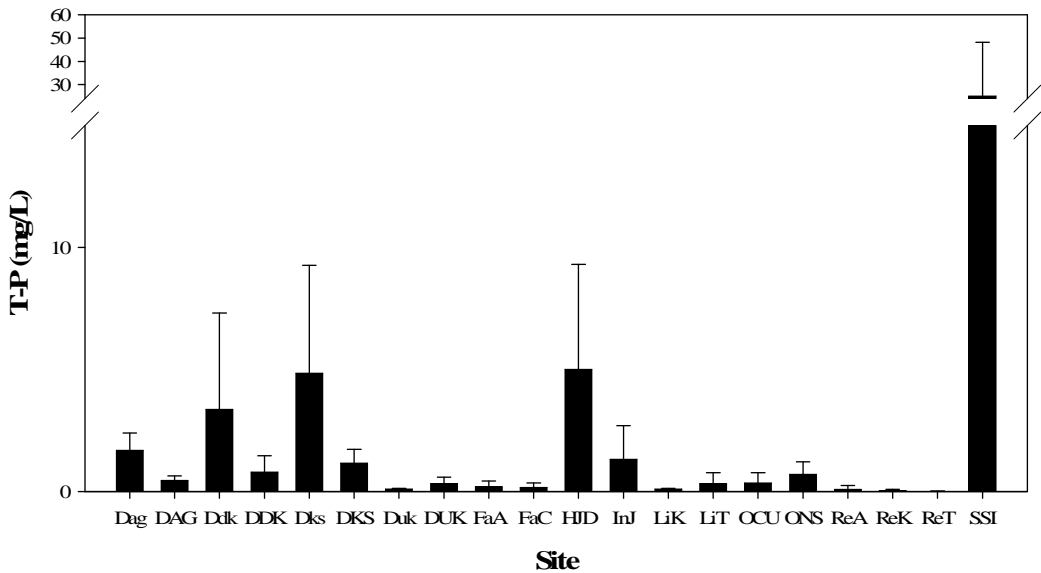


Fig. 3.11. Average total phosphorus of water samples from the sampling sites over the entire sampling period.

### 3.2.1.8. Total nitrogen

총질소(T-N)의 경우 연평균 0.96 ~ 254.69 mg/L으로 총인 오염도에 비해 더 높게 나타났다(Fig. 3.12). 특히 유기물 오염지역인 Ddk 지점과 Dks 지점, 산업폐수 오염지역인 SSI 지점이 각각 179.17 mg/L, 193.96 mg/L, 254.69mg/L로 가장 높았다. 오염지역별 현황을 보면 대조조사지점이 0.97 mg/L ~ 1.65 mg/L, 회복지인 DAG 지점이 3.18 mg/L로 전체적으로 낮은 오염도를 보였다. 농경지 오염지의 경우 FaC 지점이 1.23 mg/L으로 대조조사지와 비슷한 수준의 오염도를 보였으나 FaA 지점은 8.32 mg/L로 비교적 높은 오염도를 나타내었다. 각 지점에 대한 계절별 변화추이를 살펴보면 대체적으로 여름철에 높은 농도를 나타내었다.

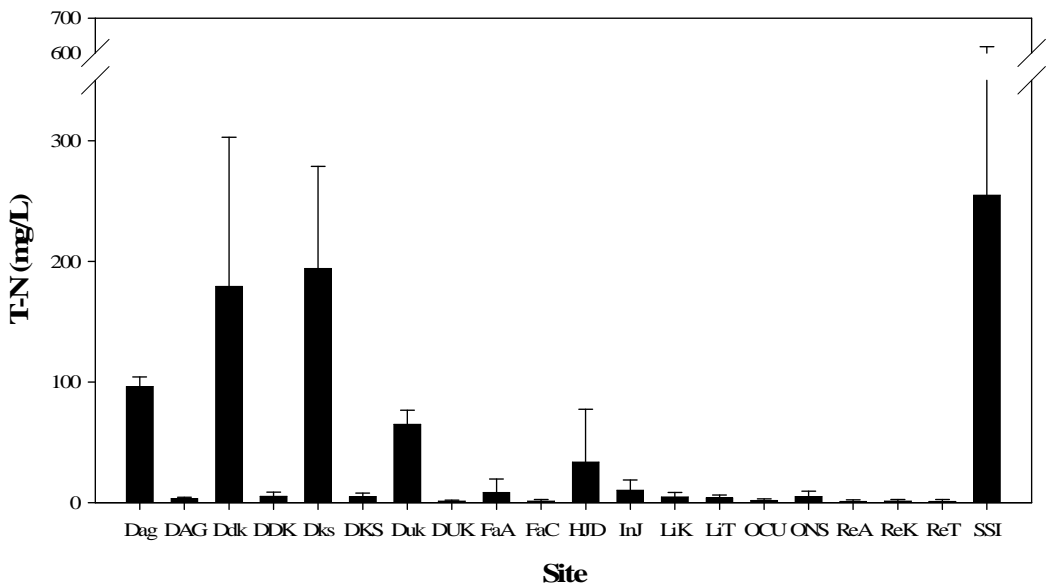


Fig. 3.12. Average total nitrogen of water samples from the sampling sites over the entire sampling period.

### 3.2.2 미소생태계의 군집구조

#### 3.2.2.1. 저서성 대형 무척추동물 군집구조

전체적으로 총 4문 6강 16목 48과 80종이 채집되었다. 오염원의 종류에 따른 군집구성을 보면 대조조사지에서 4문 6강 16목 37과 59종으로 가장 높았고, 산업폐수 오염지역이 3문 5강 9목 16과 16종으로 가장 낮았다 (Table 3.2). 산업폐수 오염지역의 경우 하루살이목과 날도래목 등에서 각각 1종만이 출현하여 다른 지역에 비해 종수가 낮게 나타났다.

전체적으로 대조조사지의 경우 다양한 분류군이 출현하였고, 경작지 오염지역과 축산폐수 오염지역에서는 하루살이목, 날도래목과 깔따구과의 상대풍부도가 높았고, 유기물 오염지역에서는 깔따구과와 빈모강이 높은 비율로 나타났다. 특히 산업폐수 오염지역에서는 오염에 내성이 강한 빈모강이 우점하여 나타났다. 그리고 강우가 많은 여름 조사(Jun 03, May 04)에서는 출현 분류군이 감소하고, 깔따구과와 빈모강이 상대적으로 높은 밀도로 출현하였다. 전체적으로 오염종류와 계절에 따른 변동을 잘 반영하였다.

군집구조의 경우 본 연구에서 조사한 대조조사지와 오염지역에서 교란정도에 따른 차이를 보였다. 대조조사지의 경우 전반적으로 다양한 생물군이 출현하는 경향을 보였으나, 유기물이나 산업폐수 오염지역의 경우 종풍부도가 감소하였으며 빈모류가 80%이상 우점하여 나타났다(Fig. 3.13). 경작지 오염지역의 경우 하루살이류와 복족류의 상대풍부도가 높았고, 축산폐수 오염지역에서는 날도래류와 깔따구류가 상대적으로 많이 출현하였다.



Table 3.2. Community composition of macroinvertebrates depending on the pollution sources\*.

| <i>Pollution source</i> | <i>Phylum</i> | <i>Class</i> | <i>Order</i> | <i>Family</i> | <i>Species</i> |
|-------------------------|---------------|--------------|--------------|---------------|----------------|
| Reference               | 4             | 6            | 16           | 37            | 59             |
| Domestic                | 4             | 5            | 10           | 20            | 23             |
| Farmin                  | 4             | 6            | 11           | 24            | 35             |
| Livestock               | 4             | 6            | 13           | 24            | 33             |
| Industry                | 3             | 5            | 9            | 16            | 16             |

\*Period: March, 2003 through December, 2004 (7 seasons).

Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).

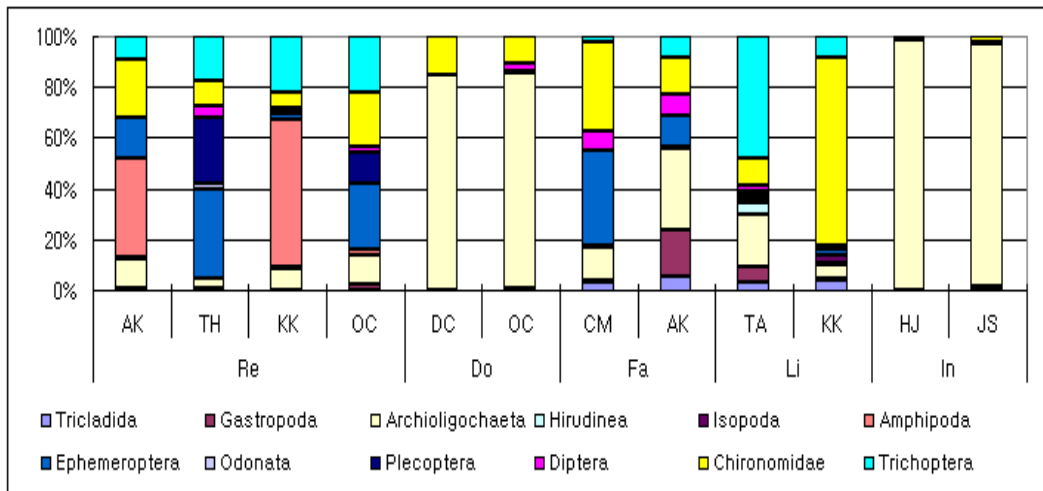


Fig. 3.13. Relative abundance of the macroinvertebrate community from the pollution sources. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).

### 3.2.2.2. 부착 및 부유 조류 군집구조

2003년 3월부터 2004년 6월까지 출현한 부착조류 분류군의 상대풍부도를 나타내었다(Fig. 3.14). 대조조사지에서는 지점에 따라 규조강(Bacillariophyceae), 남조강(Cyanophyceae)과 와편모조강(Dinophyceae)의 상대풍부도가 높았다. 유기물과 경작지오염지역에서도 주로 규조강의 출현비율이 높았고, 축산폐수 오염지역인 LiK의 경우 녹조강(Chlorophyceae)이 높게 출현하였다. 산업폐수 오염지역인 HJD의 경우 남조강이 78%이상 출현하였다.

2004년 9월과 12월 조사에서 규조류가 83종, 녹조류가 26종, 남조류가 12종, 기타 4종이 출현하였다(Fig. 3.15). 오염종류에 따라서도 규조류의 종수가 가장 많았다. 그 다음으로 녹조류, 남조류가 출현하였다. 축산폐수오염지역에서 규조류와 녹조류의 종수가 높게 나타났고, 산업폐수오염지역에서 규조류의 종수가 가장 낮았다.

Table 3.3은 2004년 9월과 12월 각 조사지점에서 조사된 부착조류의 종수, 다양도지수, DAIPo 지수와 총밀도를 나타내었다. 전체적으로 12월 축산폐수오염지역인 Th에서 35종으로 가장 많은 종이 출현하였고, 12월 산업폐수 오염지역인 HJD에서 5종으로 가장 낮았다. 다양도지수의 경우에도 12월 축산폐수 오염지역인 Th에서 3.77로 가장 높았고, 9월 산업폐수 오염지역인 InJ에서 0.53으로 가장 낮았다. DAIPo지수의 경우 50이상이면  $\beta$ -중부수성을 나타낸다. 9월 조사에서는 대조조사지, 경작지와 축산폐수 오염지역에서는 DAIPo 지수가 50이상으로 나타났고, 유기물오염지역은 평균 30정도로  $\alpha$ -중부수성을 나타내었고, 특히 산업폐수 오염지역은 10이하로 polysaprobity를 나타내면 매우 오염이 심함을 보여주었다. 12월 조사에서는 대부분의 지점에서 DAIPo지수가 50이상으로 나타났다. 전체밀도는 9월 조사에서 산업폐수오염지역에서 가장 높게 나타났고, 대조조사지에서 총밀도가 가장 낮았다.

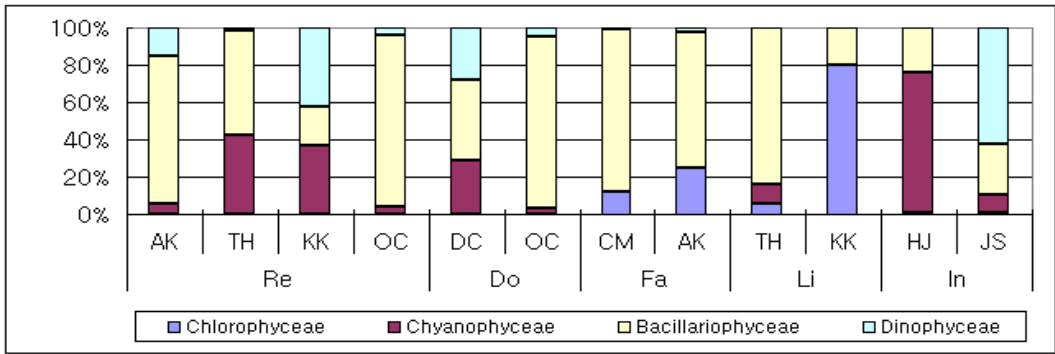


Fig. 3.14. Relative abundance of the attached algae community from the pollution sources. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).

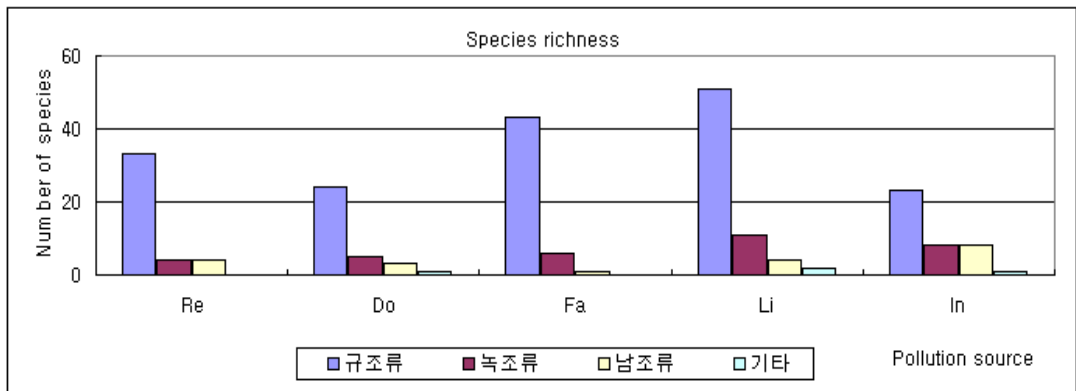


Fig. 3.15. Number of the attached algae species from the pollution sources. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).

Table 3.3. Biotic indices of the attached algae depending on pollution source.

| <i>Date</i>        | <i>Pollution source</i> | <i>Sampling site</i> | <i>SR*</i> | <i>Diversity</i> | <i>DAIpo**</i> | <i>Density (cells/ml)</i> |
|--------------------|-------------------------|----------------------|------------|------------------|----------------|---------------------------|
| September,<br>2004 | Reference               | ReA                  | 6          | 2.22             | 40.00          | 3,645                     |
|                    |                         | ReT                  | 6          | 2.13             | 84.85          | 1,980                     |
|                    |                         | ReK                  | 9          | 2.57             | 63.35          | 9,315                     |
|                    |                         | OCU                  | 6          | 2.13             | 75.03          | 2,574                     |
|                    | Domestic<br>sewage      | DDK                  | 13         | 1.79             | 27.12          | 248,580                   |
|                    |                         | ONS                  | 14         | 2.01             | 43.28          | 19,260                    |
|                    | Farming                 | FaC                  | 17         | 3.70             | 82.08          | 57,600                    |
|                    |                         | FaA                  | 17         | 2.09             | 75.00          | 347,400                   |
|                    | Livestock               | LiT                  | 19         | 2.22             | 64.09          | 113,850                   |
|                    |                         | LiK                  | 21         | 3.21             | 58.81          | 263,430                   |
|                    | Industry                | HJD                  | 10         | 1.87             | 5.38           | 4,513,500                 |
|                    |                         | InJ                  | 15         | 0.53             | 5.01           | 1,678,725                 |
| December,<br>2004  | Reference               | ReA                  | 14         | 1.40             | 84.95          | 64,620                    |
|                    |                         | ReT                  | 9          | 1.97             | 99.62          | 28,845                    |
|                    |                         | ReK                  | 7          | 2.29             | 87.80          | 4,260                     |
|                    |                         | OCU                  | 9          | 1.11             | 54.75          | 17,739                    |
|                    | Domestic<br>sewage      | DDK                  | 7          | 2.30             | 77.39          | 597,000                   |
|                    |                         | ONS                  | 19         | 1.94             | 85.04          | 528,300                   |
|                    | Farming                 | FaC                  | 28         | 3.18             | 51.17          | 1,001,100                 |
|                    |                         | FaA                  | 26         | 2.54             | 67.74          | 259,245                   |
|                    | Livestock               | LiT                  | 35         | 3.77             | 86.32          | 264,600                   |
|                    |                         | LiK                  | 19         | 2.15             | 78.60          | 498,780                   |
|                    | Industry                | HJD                  | 25         | 3.34             | 53.04          | 549,450                   |
|                    |                         | InJ                  | 5          | 1.73             | 21.43          | 2,730                     |

\* SR: Species richness.

\*\*DAIpo: Diatom assemblage index of organic water pollution.

### 3.2.2.3. 미생물 군집구조

채집한 샘플로부터 각 지점별 미생물(진정 세균계)의 군집 구조를 PCR-DGGE 분석 기법으로 분석하였다. 오염하천의 저니 미생물 군집을 조사하기 위하여 직접 저니에서 추출한 DNA를 이용하여 primer로 GC-341F와 518R을 사용하여 touchdown-PCR을 수행하여 얻은 증폭산물을 denaturing gradient gel에 loading 한 결과는 Fig. 3.16에 나타내었다.

전반적으로 청정지역이 다양성이 높은 군집을 보여주는 DGGE profile을 볼 수 있었다. 그리고 이들 band의 profile은 서로 상이한 유형을 보였다. 오염지역에 큰 관계없이 유사한 몇 가지 우점종 profile을 보였다. 또한 유기물 오염지와 산업폐수 오염지역에서는 대체적으로 band수도 적었으며 다양한 profile을 볼 수 없었다. 이는 상류지역(비오염지역)에 비하여, 다양한 유기물질이 유입되는 중류이후의 지점에서 음식점 업소오수 및 생활오수의 유입으로 인한 비슷한 유기물질이 유입되어 미생물 군집이 오염상황에 적응된 때문인 것으로 사료된다. 또한 오염 특성에 따라 특이적인 우점종 또는 특이종 profile이 관찰되었다. profile의 전체적인 특성은 채집시기에 따라 영향을 많이 받는 것으로 나타났다. 전반적으로 *Acinetobacter* sp., *Sphingomonas* sp., *Clostridium* sp., *Propionibacterium* sp., *Exiguobacterium* sp. 및 *Pseudomonas* sp.과 같은 종들이 다양한 지역에 걸쳐 우점종으로 나타났으며, 나열된 순으로 가장 우점하는 것으로 나타났다.

각 지점에서의 미생물 군집구조를 분석하기 위해 PCR-DGGE gel 상의 이미지의 pattern을 분석하여 덴드로그램으로 나타내었다(Fig. 3.17). 각 지점에 대한 종풍부도를 이용하여 cluster 한 결과 크게 3그룹(I, II 및 III)으로 분류되었다. I 그룹은 2002년 1월부터 2004년 1월까지 채집 분석된 Duk, Ddk, Dks, Dkh 및 Dag 지점이 분류되었고, II 그룹은 2003년 3월부터 2004년 12월까지 채집 분석된 ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaK, LiT, LiK, HJD 및 InJ 지점이 분류되었으며, III 그룹은 2005년 2월부터 2007년 4월까지 채집 분석된 DUK, DDK, DKS, DAG, OCU, ONS, SSI 및 HJD 지점이 분류되었다. 그리고 각 그룹 내에서도 3개의 그룹으로 나뉘어졌는데 대체적으로 오염도에 따라 지점이 묶이는 것으로 나타났다. 이는 각 지점에 대한 시료채취 시기가 다르기 때문인 것으로 사료된다.

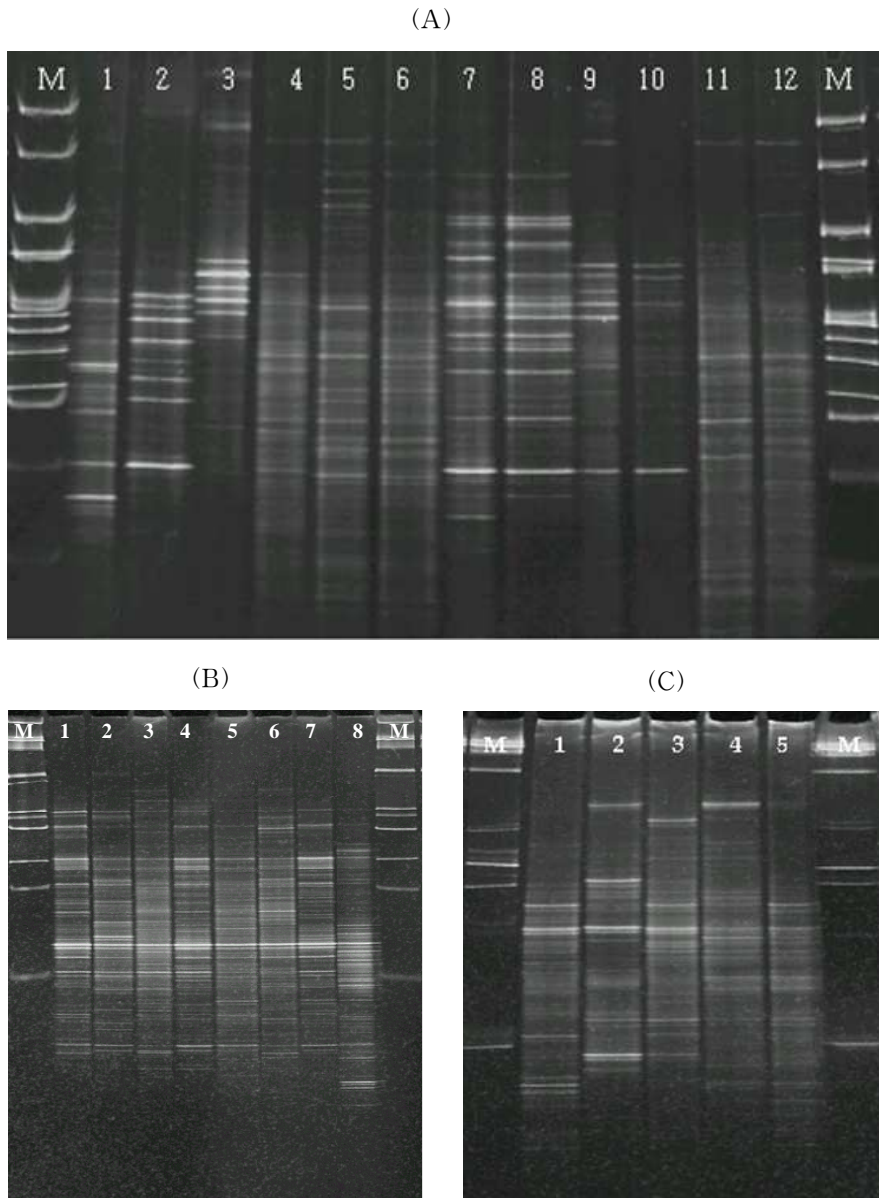


Fig. 3.16. PCR-DGGE profile of microbial communities. (A): February, 2004. Lane 1: ReA; 2: ReT; 3: ReK; 4: OCU; 5: DDK; 6: ONS; 7: FaC; 8: FaA; 9: LiT; 10: LiK; 11: HJD; 12: InJ; M: marker. (B): February, 2007. Lane 1: DUK; 2: DDK; 13: DKS; 4: DAG; 5: OCU; 6: ONS; 7: HJD; 8: SSI; M: marker. (C): February, 2003. Lane 1: Duk; 2: Dkh; 3: Ddk; 4: Dks; 5: Dag; M: marker.

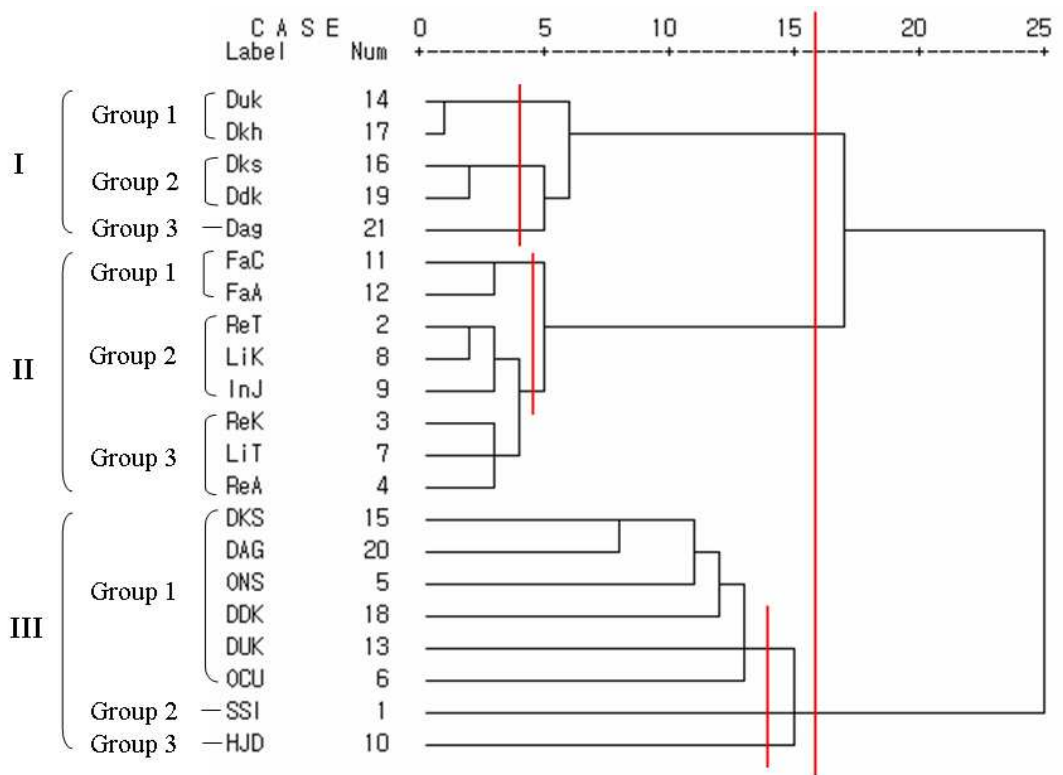


Fig. 3.17. Cluster analysis of microbial communities based on species richness over the entire sampling period..

Table 3.4. Identification of selected PCR-DGGE bands based on 16S rDNA sequence.

| Microorganisms                           | Identity       | Access No. | Sites to be found specifically  |
|--|----------------|------------|---------------------------------|
| <i>Acidobacterium capsulatum</i>         | 172/186 (92%)  | X97099     |                                 |
| <i>Acidovorax</i> sp.                    | 194/194 (100%) | AY788961   |                                 |
| <i>Acinetobacter lwoffii</i>             | 195/195 (100%) | AY136083   |                                 |
| <i>Acinetobacter</i> sp.                 | 195/195 (100%) | AY786323   |                                 |
| <i>Acinetobacter</i> sp.                 | 195/195 (100%) | AJ301674   |                                 |
| <i>Acinetobacter</i> sp.                 | 197/197 (100%) | AY043369   |                                 |
| <i>Acinetobacter</i> sp.                 | 193/195 (98%)  | AY786323   |                                 |
| Agricultural soil bacterium              | 184/194 (94%)  | AJ252618   |                                 |
| <i>Arthrobacter</i> sp.                  | 174/174 (100%) | AY437617   |                                 |
| <i>Bacillus luciferensis</i>             | 194/195 (99%)  | AJ419629   |                                 |
| <i>Bacillus psychrodurans</i>            | 193/194 (99%)  | BPS277984  |                                 |
| <i>Bacillus</i> sp.                      | 194/194 (100%) | BSP431331  |                                 |
| <i>Bacillus</i> sp.                      | 195/195 (100%) | AY269875   |                                 |
| <i>Blastomonas</i> sp.                   |                |            |                                 |
| <i>Clostridium aff. Estertheticum</i>    | 169/169 (100%) | CAF297442  | Reference site                  |
| <i>Clostridium lituseburense</i>         | 168/168 (100%) | AY458860   | Reference site                  |
| <i>Clostridium</i> sp.                   | 168/168 (100%) | AY188845   |                                 |
| <i>Clostridium</i> sp.                   | 154/154 (100%) | AY188849   |                                 |
| <i>Dermatophilus</i> like sp.            | 167/168 (99%)  | AJ244775   | Reference site                  |
| <i>Erythrobacter</i> sp.                 | 169/169 (100%) | AY437632   |                                 |
| <i>Erythrobacter</i> sp.                 | 163/164 (99%)  | AY167828   |                                 |
| <i>Exiguobacterium</i> sp.               | 193/194 (99%)  | AY831656   |                                 |
| <i>Exiguobacterium</i> sp.               |                |            |                                 |
| <i>Flavobacteria</i> bacterium           | 184/188 (97%)  | DQ004245   |                                 |
| Glacier bacterium                        | 186/194 (95%)  | AY315180   | Farming site<br>Livestock site  |
| <i>Luteibactor rhizovicina</i>           | 189/194 (97%)  | AJ580499   | Reference site<br>Faming site   |
| <i>Methylobacterium hispanicum</i>       | 166/167 (99%)  | AJ785570   | Reference site                  |
| <i>Moraxella canis</i>                   | 177/194 (91%)  | AJ269511   | Reference site                  |
| <i>Pandoraea</i> sp.                     | 186/186 (100%) | AY268171   |                                 |
| Phenanthrene-degrading bacterium         | 169/169 (100%) | AY177363   |                                 |
| <i>Photorhabdus luminescens</i>          | 192/193 (99%)  | AY444555   |                                 |
| <i>Pseudaminobacter salicylatoxidans</i> | 164/169 (97%)  | AF072542   | Livestock site<br>Industry site |
| <i>Pseudomonas fluorescens</i>           | 194/194 (100%) | AY321588   |                                 |



(Table 3.4. continued)

| Microorganisms                              | Identity       | Access No. | Sites to be found specifically  |
|---|----------------|------------|---------------------------------|
| <i>Pseudomonas</i> sp.                      | 193/194 (99%)  | AY456712   |                                 |
| <i>Pseudomonas</i> sp.                      | 194/194 (100%) | AF332541   |                                 |
| <i>Pseudoxanthomonas taiwanensis</i>        | 190/194 (97%)  | AB210278   | Domestic sewage<br>Farming site |
| <i>Ralstonia</i> sp.                        | 194/194 (100%) | AY177368   |                                 |
| <i>Ruminococcus palustris</i>               | 194/194 (100%) | AJ296179   | Reference site                  |
| <i>Sphingomonas</i> sp.                     | 171/171 (100%) | AY167827   |                                 |
| <i>Sphingomonas</i> sp.                     | 167/169 (98%)  | AJ617690   |                                 |
| <i>Sphingomonas</i> sp.                     | 169/169 (100%) | U37347     |                                 |
| <i>Sphingomonas</i> sp.                     | 166/169 (98%)  | AF128870   |                                 |
| <i>Sphingomonas</i> sp.                     | 169/169 (100%) | AJ575818   |                                 |
| <i>Sphingomonas</i> sp.                     | 169/169 (100%) | AY734544   |                                 |
| <i>Stenotrophomonas maltophilia</i>         | 186/186 (100%) | AY445079   |                                 |
| <i>Thiothrix</i> sp.                        | 157/157 (100%) | AF148516   |                                 |
| <i>Trichococcus flocculiformis</i>          | 194/194 (100%) | AJ306611   |                                 |
| Uncultured <i>Acidobacteria</i> bacterium   | 182/182 (100%) | AJ582044   |                                 |
| Uncultured <i>Acidobacteria</i> bacterium   | 169/169 (100%) | AY217488   |                                 |
| Uncultured <i>Acinobacterium</i>            |                |            | Reference site                  |
| Uncultured bacterium                        | 193/194 (99%)  | AY570568   |                                 |
| Uncultured <i>Bacteroides</i> sp.           |                |            |                                 |
| Uncultured <i>Clostridia</i> bacterium      | 165/169 (97%)  | AY607225   |                                 |
| Uncultured <i>cyanobacterium</i>            | 171/172 (99%)  | AY710671   |                                 |
| Uncultured freshwater bacterium             | 195/195 (100%) | DQ065476   |                                 |
| Uncultured gammaproteobacterium             | 194/194 (100%) | AJ888552   |                                 |
| Uncultured <i>Legionella</i> sp.            | 190/194 (97%)  | AY924132   |                                 |
| Uncultured <i>Meiothermus</i> sp.           |                |            |                                 |
| Uncultured <i>Nirosospira</i> sp.           |                |            | Industry site                   |
| Uncultured <i>Propionicimonas</i> sp.       | 161/174 (92%)  | AY675983   | Reference site                  |
| Uncultured soil bacterium                   | 186/186 (100%) | AY734276   |                                 |
| Uncultured <i>Verrucomicrobia</i> bacterium | 176/184 (95%)  | AY043933   |                                 |
| Uncultured <i>Xanthomonadaceae</i> sp.      |                |            |                                 |

### 3.3. 결론

본 연구에서는 하천 생태계 관리를 위한 이화학·생물학적 통합수질지표 개발을 위한 기초자료를 수집하기 위해 미소생태계 내의 환경요인과 생물요인을 분석하였고 이를 통해 다음과 같은 결론을 얻을 수 있었다.

1. 생물군집(미생물, 조류 및 저서생물) 분석결과 이화학적 환경자료와 각 생물군의 군집들은 대체로 이들 지역의 특성에 따른 군집변화가 이루어짐을 관찰하였다. 특히 저서생물의 경우 청정지역과 오염지역 종류들 간에 뚜렷이 구분되는 양상을 보였다. 따라서 조사지점의 선택은 적절하였고 재현성 있는 결과를 거두는데 기여하였다고 생각된다.
2. 환경요인 및 생물요인 분석결과 특정지역에 따라 계절적인 영향을 많이 받는 것으로 관찰되었다. 환경요인의 경우는 계절별 강우량의 차이에 의해 영향을 받기도 하지만 인간 활동영역에 근접하게 위치한 지점일수록 더 영향을 받는 것으로 관찰되었다.
3. 오염도에 따른 분류군별 군집구조를 보면 대형무척추동물과 조류는 오염도에 따라 출현하는 종들 또는 특정 종들이 뚜렷이 관찰되었다. 그러나 미생물의 경우 분석상의 오차로 인해 몇몇 종들을 제외하고는 특성화하기가 어려웠다. 이를 극복하기 위해서는 DGGE gel 간의 직접적 비교에 있어서 internal standard DNA를 첨가하는 등의 normalization 기법을 도입함으로써 data 해석상의 신빙성을 제고할 수 있을 것으로 본다.
4. 환경요인 및 생물요인 분석에 있어 일반적으로 수질평가에 이용되는 항목들에 대해 조사를 수행하였다. 통합수질지수를 개발함에 있어 환경과 생물과의 연관성을 조사하는 것은 필수적이다. 그러나 일반적인 수질항목들 만으로 생물종들의 특성을 대별하기는 어렵다. 추후 연구에서는 조사지역을 좀 더 다양하게 하고, 다양한 항목을 측정하여 수질과의 상관성을 높일 필요가 있다.

## 제 4장 통합지수개발을 위한 사전 연구

환경 및 생물요인 분석을 통해 수집된 복잡한 생태조사 자료를 분류 및 유형화를 위해 self-organizing mapping 기법을 이용하였다. 이를 통해 군집의 유형(오염상태)을 판별할 수 있다. 본 연구에서는 오염종류에 따라 분석한 분류군의 군집을 집중적으로 유형화 하고자 하였다.

또한 수질지수를 통합적으로 구현하기 위한 기초 작업으로 조사된 군집자료와 주요 환경요인과의 관련성을 검토하였다. 이를 위해 군집-환경 인과성을 찾을 수 있는 인공신경망인 다층퍼셉트론(multi layer perceptron, MLP)을 이용하였다. 또한 환경요인에 대한 각 분류군의 반응을 보기위해 학습된 다층퍼셉트론을 저서생물, 미생물 및 부착조류 군집에 적용하여 민감도를 분석하여 환경요인과 다분류군 군집구조간의 상관관계를 파악하고자 하였다.

통합수질지수 개발을 위해 생태자료를 통합하기 전에 각 분류군 생물지수의 연관성을 조사하였다. 이를 위해 각 분류군의 대표적인 생물지수들을 선정하여 인공신경망인 SOM을 이용하여 유형화 하였다. 각 분류군에서 이용된 지수는 다음과 같다. 사전에 생물지수 연구가 많이 수행되어진 저서성 대형무척추동물은 종풍부도(SRben), 총밀도(TDben), EPT 풍부도(EPTben), EPT 밀도(EPT-Dben), BMWP 지수(BMWPben), 다양도지수(Hben)를 검토하였다. 부착조류의 경우 종풍부도(SRalg)와 다양도지수(Halg) 및 총밀도(TDalg)를, 미생물의 경우 종풍부도(SRmic), 총밀도(TDmic)와 다양도지수(Hmic)를 적용하였다.

## 4.1. 연구배경

통합수질지수 개발을 위해 생태자료를 통합하기 전에 각 분류군 생물지수의 연관성을 조사하였다. 이를 위해 각 분류군의 대표적인 생물지수들을 선정하여 인공지능망인 SOM을 이용하여 유형화 하였다. 복잡한 생태 자료의 분류를 위해 비지도학습(non-supervised) 인공지능회로망의 방법 중 Kohonen network를 이용한 self-organizing map(Kohonen, 1989)이 가장 널리 사용되고 있다. 이 계산법은 복잡한 자료에 대하여 사전지식이 없어도 data mining을 효율적으로 수행할 수 있다. 수질이나 생태계조사에 있어 SOM을 이용함으로써 군집 각각의 유형을 판별(오염상태 등)(Chon, et al., 1996), 오염종류에 따라 채집되는 군집의 유형화 그리고 중요 분류군간의 연계성을 유형화 할 수 있다 (legendre and legendre, 1983; Kim et al., 2008).

또한 다층 퍼셉트론(Multi-layer perceptron)은 목표 값을 설정해 주어 그 값에 가까운 출력 값을 내도록 학습시키는 지도학습알고리즘의 하나로 복잡한 자료의 연관 유형성을 파악하는데 뛰어나다. 입력층에는 서식환경(예, 유속, 하상구성, 유기물량 등)을, 출력층에는 이에 대응되는 각종 분류군(저서성 대형무척추동물, 미생물, 조류 등의 주요 지표군)을 입력시켜 주어진 서식처 환경에 따라 생물의 서식 연관성을 유형화 할 수 있다. 이 후 별도의 새로운 자료를 주어 이에 해당되는 환경 여건 또는 조건에 따라 어느 정도로 생물이 서식하거나 지수를 나타내는지 예측할 수 있다(Chon, et al., 1996).

## 4.2. 실험방법

### 4.2.1. Self-organizing mapping(SOM)을 이용한 군집 유형화

본 연구에서는 복잡한 생태 자료의 분류를 위해서 Kohonen network을 이용한 self-organizing map(Kohonen, 1984)을 이용하였다(Fig. 4.1) (Chon, 1996). 입력에서 binary가 아닌 연속변량을 허용하여 개체군 크기 등 연속적으로 계측되는 수치를 사용할 수 있고, 비지도학습(unsupervised learning)을 통해 자기조직적으로 입력정보의 유형결정을 할 수 있는 특징이 있다.

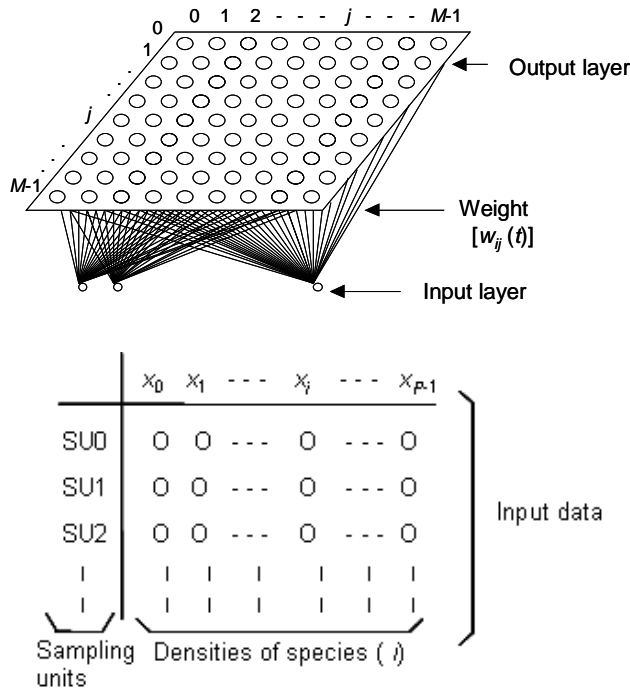


Fig. 4.1. Schematic diagram of the Kohonen network (adapted from Chon et al., 1996; Kwak et al., 2000)

각각의 군집(미생물, 저서생물, 부착조류)에 대해 SOM을 적용하였다. 저서성 무척추동물의 경우 중요 분류군(예; 파리류, 하루살이류, 날도래류, 지렁이류)에 따라 종 수와 출현개체군으로 나누어 밀도를 나타내고 이를 자연로그 변환한 후 0.01 ~ 0.99 범위에서 종 수와 개체수 각각에 대해 정규화 하였다.

중요 분류군의 밀도(개체수/m<sup>2</sup>)를 나타내는 계산단위 벡터,  $\vec{x}_i$ 에 대해 아래와 같이 입력강도(input intensity),  $I_i$ 를 계산하였다.

$$\begin{aligned}
 I_i &= D(\vec{w}_i, \vec{x}) \\
 \vec{w} &= \text{가중치}, (w_{i1}, w_{i2}, \dots, w_{in})^T \\
 \vec{x} &= \text{입력벡터}, (x_1, x_2, \dots, x_n)^T \\
 D(\vec{w}_i, \vec{x}) &= \text{거리측정함수}
 \end{aligned}$$

출력뉴런  $j$ 와 입력뉴런  $i$ 는 연결계수(weight)  $w_{ji}$ 로 서로 연결이 되어 있으며, 연결계수는 일련의 계산 반복과정을 거쳐 적응적으로 변화하게 되는데 이 과정을 학습이라 한다. 연결계수는 초기에는 무작위적으로 작은 수로 설정이 된다. 입력벡터( $x$ )가 주어지면 출력층 뉴런( $j$ ) 각각에 대하여 입력값( $x_i$ ; 분류군  $i$ 의 밀도)과 연결계수( $w_{ji}$ )의 차이의 합( $d_i(t)$ )이 식 4.1과 같이 계산된다 (Kohonen, 1984).

$$d_j(t) = \sum_{i=0}^{N-1} (x_i - w_{ij}(t))^2 \quad (4.1)$$

계산된 값  $d_i(t)$  중에서 가장 작은 값을 가진 뉴런  $j$ 가 승자뉴런(winning neuron)으로 선택된다. 이후 가장 작은 입력강도를 가진 계산단위(또는 변량),  $x_i$ 의  $z_i$ 치를 1로 하고 다른 단위들의  $z_i$ 치는 전부 0으로 하는 경쟁과정을 거친다. 다음 단계로 이긴 단위에는 아래와 같이 가중치가 수정된다.

$$\vec{w}_i^{*w} = (1 - a(t)) \vec{w}_i^{old} + a(t) \vec{x}$$

이가지 못하는 계산단위들은 아래와 같이 된다.

$$\vec{w}_i^{*w} = \vec{w}_i^{old}$$

여기서  $0 < a < 1$ 인  $a(t)$ 는 학습계수(learning coefficient)이다. 대개 0.8에서 시작하여 점차적으로 작아지는데, 적절한 계획과정(scheduling process)

(Hecht-Nielsen, 1990)에 의해 진행될 수 있다. 훈련(training)이 끝나면 벡터는 계산단위  $n$ 개 중의  $Z_i$  요소를 선정하는데 있어서 균확률성(equiprobability)이 된다. Kohonen network에 대한 자세한 방법은 Kohonen (1984), Lipmann (1987), Hecht-Nielsen (1990), Chon et al. (1996) 등에 언급되어 있다.

#### 4.2.2. 다층퍼셉트론을 이용한 생물군집과 환경요인의 연관성

다층 퍼셉트론(Multi-layer perceptron)은 목표 값을 설정해 주어 그 값에 가까운 출력 값을 내도록 학습시키는 지도학습알고리즘의 하나로 복잡한 자료의 연관 유형성을 파악하는데 유용하다. 다층 퍼셉트론은 입력층, 은닉층, 출력층으로 구성되어 있다(Lipmann, 1987)(Fig. 4.2). 은닉층은 학습과정에서 입력과 출력사이의 관계를 결정지어주는 역할을 한다. 각 층과 층사이의 노드는 연결강도로 이어져 있는데, 학습을 하는 동안 추출된 정보가 이 연결강도에 점차적으로 담기게 된다. 출력 층에서는 각 입력 유형(pattern)에 대한 출력이 나오고 다시 이는 모본을 나타내는 목표값(target)에 가깝게 나오도록 정보를 은닉층으로 역전파하여 연결강도를 조정하도록 하며, 결과적으로 오차-목표값(error-target)과 실제 출력값 사이의 차이가 줄어들도록 한다.

목표값(target)과 실제 출력값 사이의 차이를 오차(error)라고 하며 학습은 오차의 제곱합(Least-squares-error)을 최소화하는 방향으로 진행된다. 즉 연결강도의 변화는 오차의 제곱이 가장 많이 감소하는 방향으로 변화하게 되는데 (gradient descent method) 이러한 학습방법을 수학적으로 표현하면 식 4.2 ~ 식 4.5와 같은 delta-rule과 같이 유도된다.

$$\frac{\partial E_p}{\partial w_{ij}} = -\delta_{pj} x_{pi} \quad (4.2)$$

$$O_{p,j} = \frac{1}{1 + \exp(-\lambda * T_{p,j})} \quad (4.3)$$

$$NET_{p,j} = \sum_i x_{p,i} w_{p,ji} \quad (4.4)$$

여기서  $E_p$ 는 유형  $p$ 에 대한 오차를 나타내고  $O_{p,j}$ 는 유형  $p$ 에 대한  $j$ 번째 출력 노드의 출력값이며,  $NET_{p,j}$ 는 유형  $p$ 에 대한  $j$ 번째 출력 노드의 net값,  $x_{p,i}$ 는 유형  $p$ 에 대한  $i$ 번째 노드의 출력값,  $w_{p,ji}$ 는  $j$ 번째 출력 노드와 그 앞단의  $i$ 번째 노드사이의 연결강도,  $\lambda$ 는 학습계수(learning coefficient)를 나타낸다.

$$\delta_{pj} = T_{pj} - O_{pj} \quad (4.5)$$



여기서  $T_{pj}$ 는 유형  $p$ 에 대한  $j$ 번째 출력 노드의 목표출력값을 나타낸다. 은닉층에서의 출력계산 방법도 위와 같은 방식으로 수행되며 오차는 식 4.6과 같이 정해진다.

$$E_p = \sum_j (t_{p,j} - o_{p,j}) \quad (4.6)$$

여기서 각 유형  $p$ 에 대한 각 노드의 오차합인 총 오차(Total error) 값은 식 4.7과 같이 계산된다.

$$E = \sum_p E_p \quad (4.7)$$

연결강도는 반복계산을 통하여 식 4.8과 같이 적응적으로 조절된다.

$$\begin{aligned} w_{p,ji}(t+1) &= w_{p,ji}(t) + \Delta w_{p,ji} \\ \Delta w_{p,ji} &= \eta \delta_{p,j} o_{p,i} \\ \delta_{p,j} &= (t_{p,j} - o_{p,j}) o_{p,j} (1 - o_{p,j}) \end{aligned} \quad (4.8)$$

여기서  $\Delta w_{p,ji}$ 은 유형  $p$ 에 대한  $w_{p,ji}$ 의 증가분이고,  $\eta$ 은 학습계수이다. 은닉층에서의 연결강도 조절은 출력층에서와 같으나,  $\delta$ 값은 식 4.9와 같이 결정된다.

$$\delta_{p,i} = o_{p,i} (1 - o_{p,i}) \sum_j \delta_{p,j} w_{p,ji} \quad (4.9)$$

여기서  $\delta_{p,i}$ 는  $i$ 번째 노드의 delta값이며,  $\delta_{p,j}$ 는  $i$ 번째 노드가 있는 층의 바로 다음 층의  $j$ 번째 노드의 delta값이다.

본 연구에서는 입력층에는 서식환경(예, 유속, 하상구성, 유기물량 등)을, 출력층에는 이에 대응되는 각종 분류군(저서성 대형무척추동물, 미생물, 조류 등의 주요 지표군)을 입력시키도록 한다. 이로써 주어진 서식처 환경에 따라 생물의 서식 연관성을 유형화 할 수 있다. 이 후 별도의 새로운 자료를 주어 이에 해당되는 환경 여건 또는 조건에 따라 어느 정도로 생물이 서식하거나 지수를 나타내는지 예측하도록 하였다.

### 4.2.3. 생물 지수

채취된 시료로부터 분류군 별 생물을 동정하고 계수하여 군집구조를 파악하기 위해 종 풍부도를 구하고 식 4.10 및 식 4.11과 같이 다양도(H)(Shannon 다양도, Hellawell, 1986)와 우점도(dominance index)를 측정하였다. 일반적으로 Shannon 다양도 지수가 3 이상이면 양호하고 1이하면 오염으로 교란의 영향이 큰 것을 의미한다. 우점도는 오염이 클수록 1에 가깝다.

$$H' = - \sum (n_i/N \log_2 (n_i/N)) \quad (4.10)$$

여기서  $n_i$ 는 분류군  $i$ 의 개체수이며,  $N$ 은 전체 개체수이다.

$$(Dominance\ index) = (n_1 + n_2)/N \quad (4.11)$$

여기서  $n_1$ 은 제 1 우점종 개체수를 나타내며,  $n_2$ 는 제 2 우점종 개체수이고,  $N$ 은 전체 개체수를 나타낸다.

또한 상기의 지수 외에도 저서성 대형무척추 동물의 경우 일반적으로 잘 알려진 EPT 풍부도(하루살이, 강도래, 날도래의 종 수, Resh et al., 1995)와 BMWP (Biological Monitoring of Working Party, National Water Council, 1981) 지수를 구했다. 부착조류는 일반적으로 알려진 DAIPo 지수(Watanabe, et al., 1986)를 이용하였다. 부착조류 지수는 DAIPo 지수가 상대적으로 일관되게 나타났으나, 앞서 언급한 바와 같이 연구 여건에 따라 전체적인 샘플의 종 목록을 구하기가 어려워 오염정도 및 조사자료 여건을 고려하여 기존의 DAIPo 지수를 modify 하였다. DAIPo 지수에서 사용되어진 주요 호오탁성 종과 광적응성 종을 택해 환산하였고, 각 종의 상대풍부도를 이용하여  $100 - (\text{호오탁성종의 상대풍부도 합}) - 1/2(\text{광적응성 종의 상대풍부도합})$ 으로 계산하여 지수를 구하였다.

미생물의 경우는 일반적으로 알려진 지수가 거의 없어 본 연구의 PCR-DGGE gel 상의 band intensity를 분석하여 상대적인 총 밀도로 산정하여 지수로 대응하였다.

### 4.3. 결과 및 고찰

#### 4.3.1. Self-organizing mapping(SOM)을 이용한 군집 유형화

##### 4.3.1.1. 미생물

조사기간 동안 출현한 미생물 군집 자료를 이용하여 군집 유형화를 수행하였다. Fig. 4.2는 95개 지점(19개월 × 5개 지점)에서 출현한 45종의 미생물 군집 자료를, Fig. 4.4는 86개 지점(7계절 × 12개 지점), 45종의 미생물 군집 자료를, Fig. 4.6은 80개 지점(10계절 × 8지점), 79종 중 16종의 미생물 군집 자료를 이용하여 그룹핑 하였다. 전체적으로 4~5개의 클러스터로 나뉘어졌는데 오염종류에 따른 그룹핑 보다 시기별로 그룹이 형성되었으나 Fig. 4.2의 클러스터 4의 경우 회복지인 Dag 지점이 계절에 관계없이 묶여져서 나타났다.

Fig. 4.3, Fig. 4.5 및 Fig. 4.7은 출현미생물의 SOM map 상에서의 분포 밀도를 나타내었다. 검은색(붉은색)은 밀도가 높음을 회색(파란색)으로 갈수록 밀도가 낮음을 나타낸다. Fig. 4.3에서는 *Aeromonas* sp., *Klebsiella* sp., *Lutaimonas* sp. 및 *Rhizobium* sp.과 같은 종들이 회복지 지점에서 높은 밀도로 분포하는 것으로 나타났다. Fig. 4.5에서 청정지역에서는 *Ruminococcus* sp., *Dermatophilus* like sp., *Arthrobacter* sp. 및 *Methylobacterium* sp. 등이, 오염지역에서는 *Acinetobacter* sp.가 높은 밀도로 분포하였으며, *Flavobacteria* sp.와 *Pseudomonas* sp. 종은 오염지역 및 청정지역 모두에서 분포하는 것으로 나타났다.

결과에서 보는 것처럼 미생물의 경우 시료채취 지역보다는 시기별로 grouping 되는 이유는 계절에 따른 수온의 변화로 인한 대사작용과 이에 따른 미생물 군집 구조형성에 기인하는 것으로 판단된다. 특히 미생물의 경우 환경적 요인 (유기물 및 영양 염류 농도 그리고 pH등)에 적응력이 강해 시료채취 장소에 별 상관없이 그 출현을 보이는 것으로 판단된다.

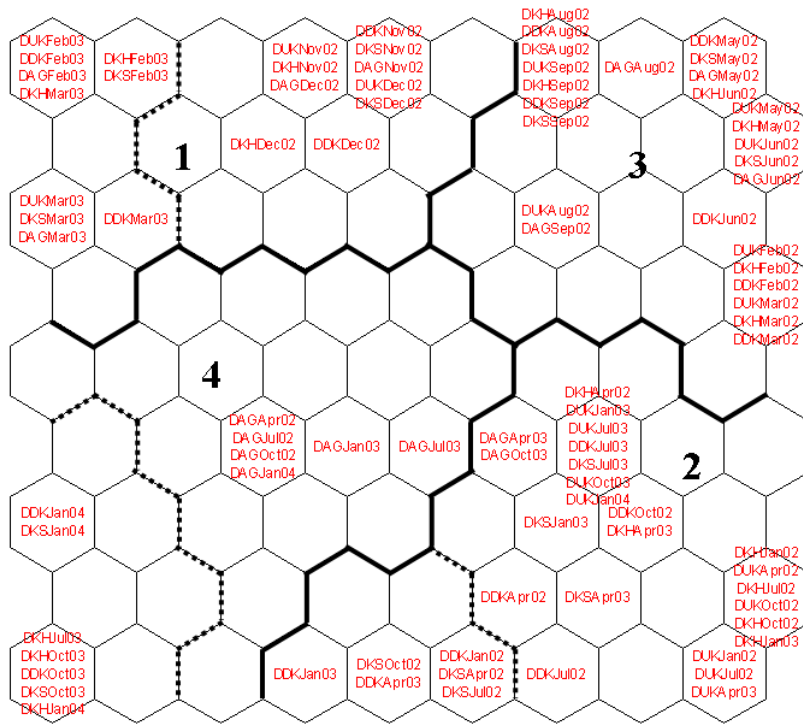


Fig. 4.2. SOM patterning of sampling sites based upon microbial density. Period: January, 2002 through January, 2004 (19 months); Sites: Duk, Ddk, Dks, Dkh, Dag (5 sites).

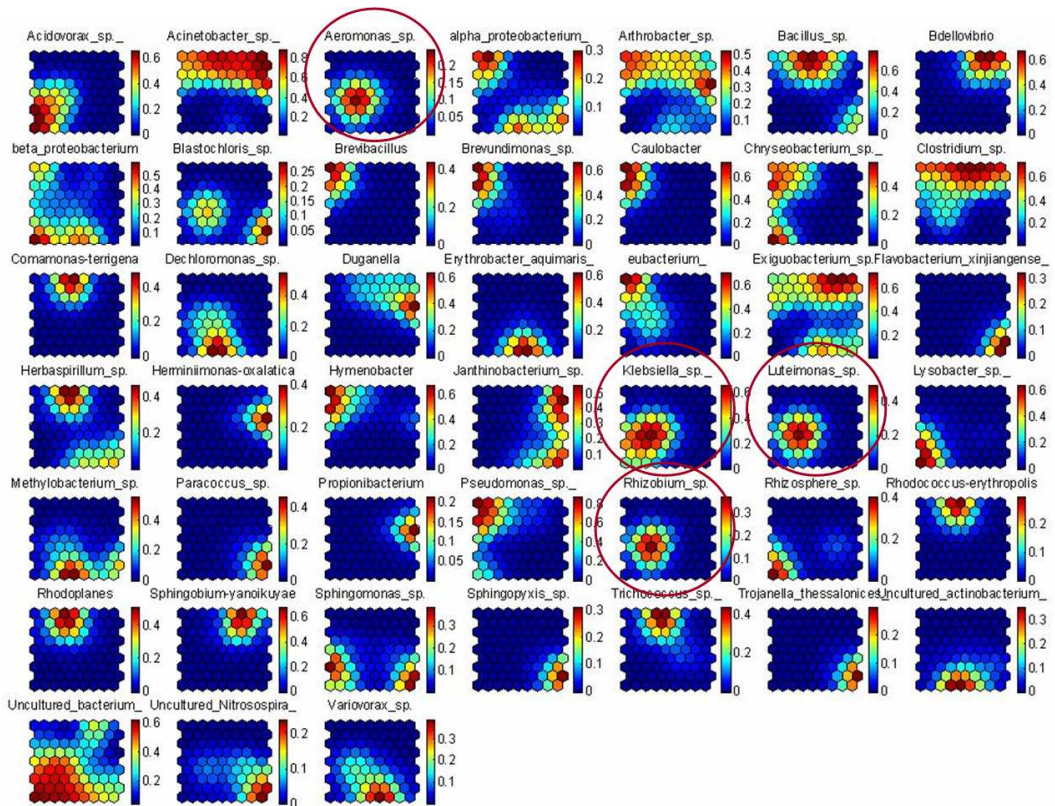


Fig. 4.3. Distribution of microbial populations on the SOM. Period: January, 2002 through January, 2004 (19 months); Sites: Duk, Ddk, Dks, Dkh, Dag (5 sites); Solid circle: restoration site.

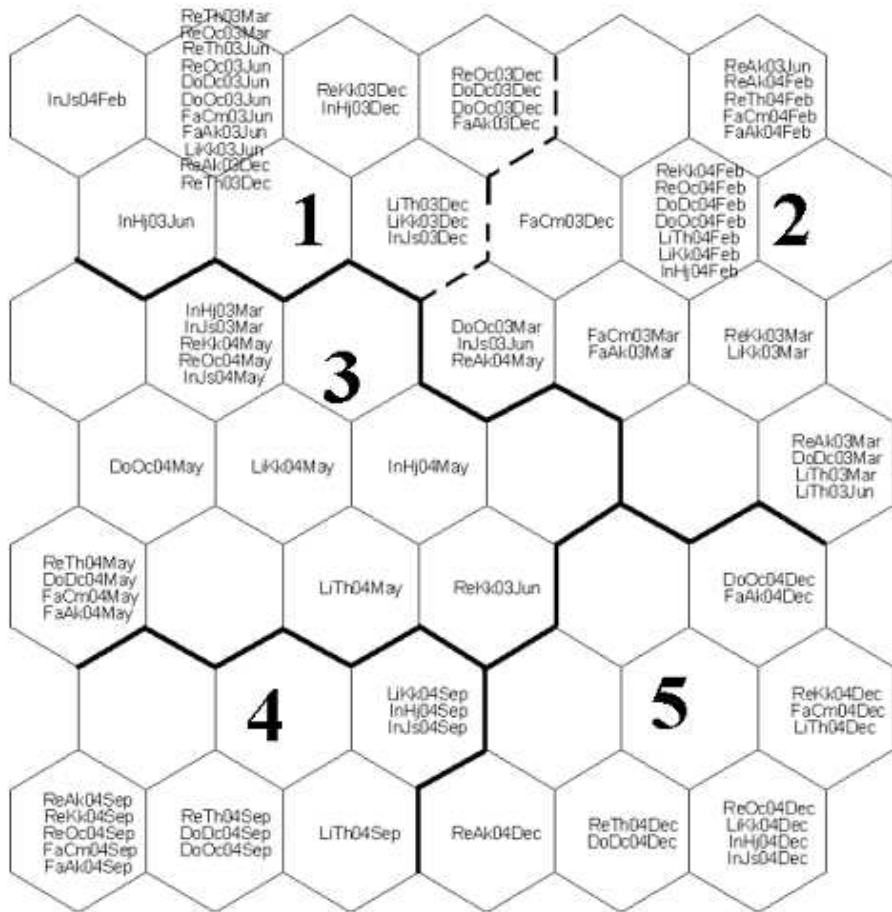


Fig. 4.4. SOM patterning of sampling sites based upon microbial density. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).

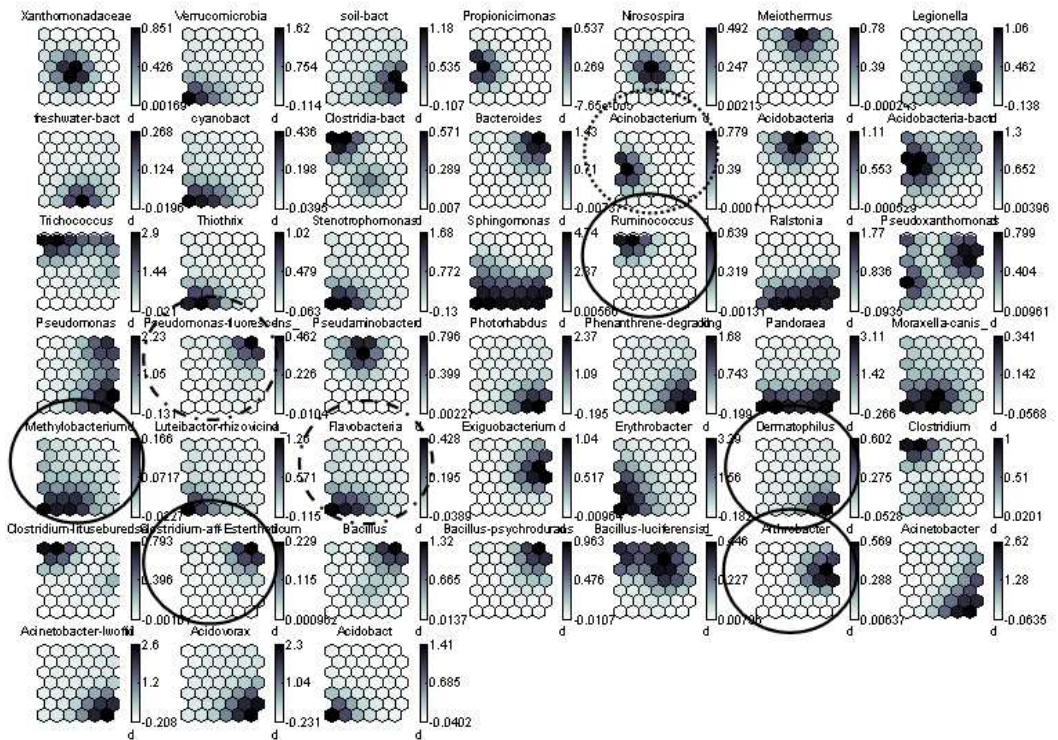


Fig. 4.5. Distribution of microbial populations on the SOM. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites). Solid circle: reference site; Dotted circle: polluted site; Broken circle: polluted and reference sites.

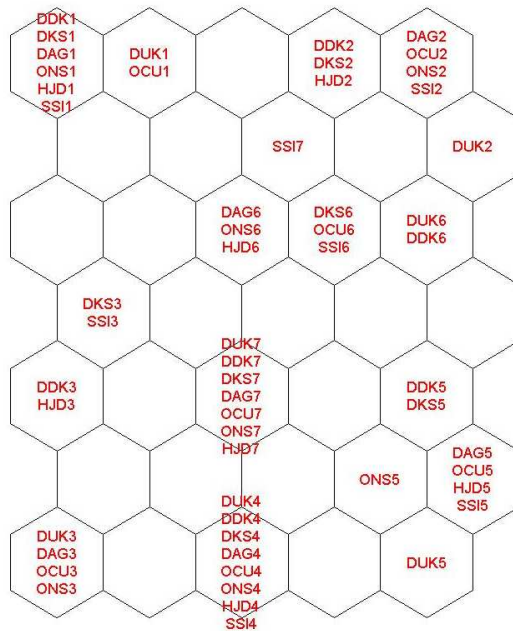


Fig. 4.6. SOM patterning of sampling sites based upon microbial density. Period: February, 2005 through April, 2007 (10 seasons); Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD (8 sites).

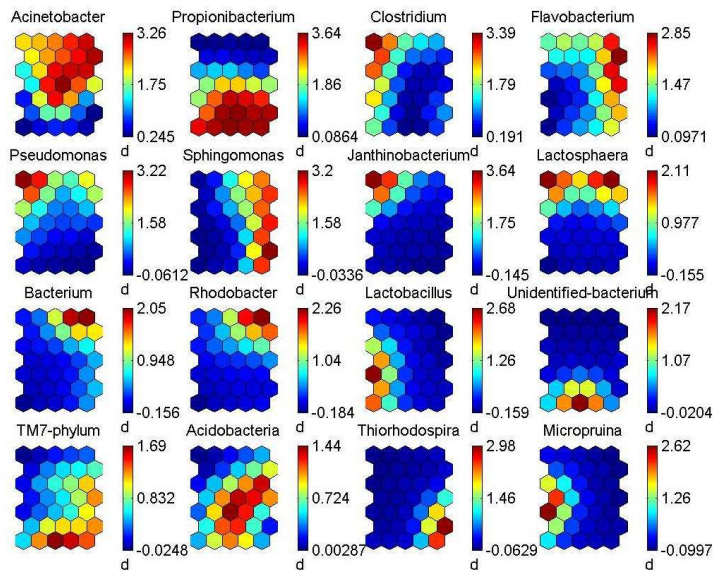


Fig. 4.7. Distribution of microbial populations on the SOM. Period: February, 2005 through April, 2007 (10 seasons); Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD (8 sites).



#### 4.3.1.2. 대형무척추동물

대형무척추동물 조사기간 및 지점은 미생물 자료와 동일하며 각 조사기간 동안의 대형무척추동물 출현량은 Fig. 4.8은 95개 지점(19개월 × 5개 지점)에서 출현한 47과의 대형무척추동물 자료를, Fig. 4.10은 86개 지점(7계절 × 12개 지점), 48과의 대형무척추동물 자료를, Fig. 4.12는 80개 지점(10계절 × 8지점), 16과의 대형무척추동물 자료를 이용하여 조사시기별 지점에 따른 그룹핑을 수행하였다. 전체적으로 5~6개의 클러스터로 나뉘어졌다. 지도의 아래인 클러스터 5와 6에는 대조조사지가 묶여져서 나타났다. Fig. 4.8에서는 클러스터 2와 3에 각각 청정지역과 회복지가 묶여졌다. Fig. 4.10에서는 지도의 위쪽인 클러스터 1과 4에서는 산업폐수와 유기물 오염지역이 묶여졌고, 클러스터 2와 3에서는 각각 경작지 오염지역과 축산폐수 오염지역이 묶여졌다. 전반적으로 오염정도에 따른 구배를 잘 보여주었다.

저서생물 자료를 이용한 SOM 지도에 나타난 각 지점에서 출현한 주요 분류군의 밀도분포를 Fig. 4.9, Fig. 4.11 및 Fig. 4.13에 나타내었다. 점선으로 표기된 원안의 분류군은 오염이 심한 상태에서 주로 출현하는 원돌이 물달팽이과(Physidae), 나방파리과(Psychodidae)와 실지렁이(Tubificidae) 등이 오염지역에서 높은 밀도로 나타났다. 실선으로 표기된 원안의 분류군은 주로 청정한 하천에서 출현하는 갈래하루살이과(Leptophlebiidae), 강도래과(Perlidae)와 우묵날도래과(Limnephilidae) 등이 청정지역 및 회복지역에서 높은 밀도로 출현하였다. 노란 원은 오염지역과 청정지역에 걸쳐져서 출현하는 꼬마하루살이과(Baetidae), 갈따구과(Chironomidae), 납작하루살이과(Heptagenidae)와 줄날도래과(Hydropsychidae)로 전체적으로 고루게 분포하였다.

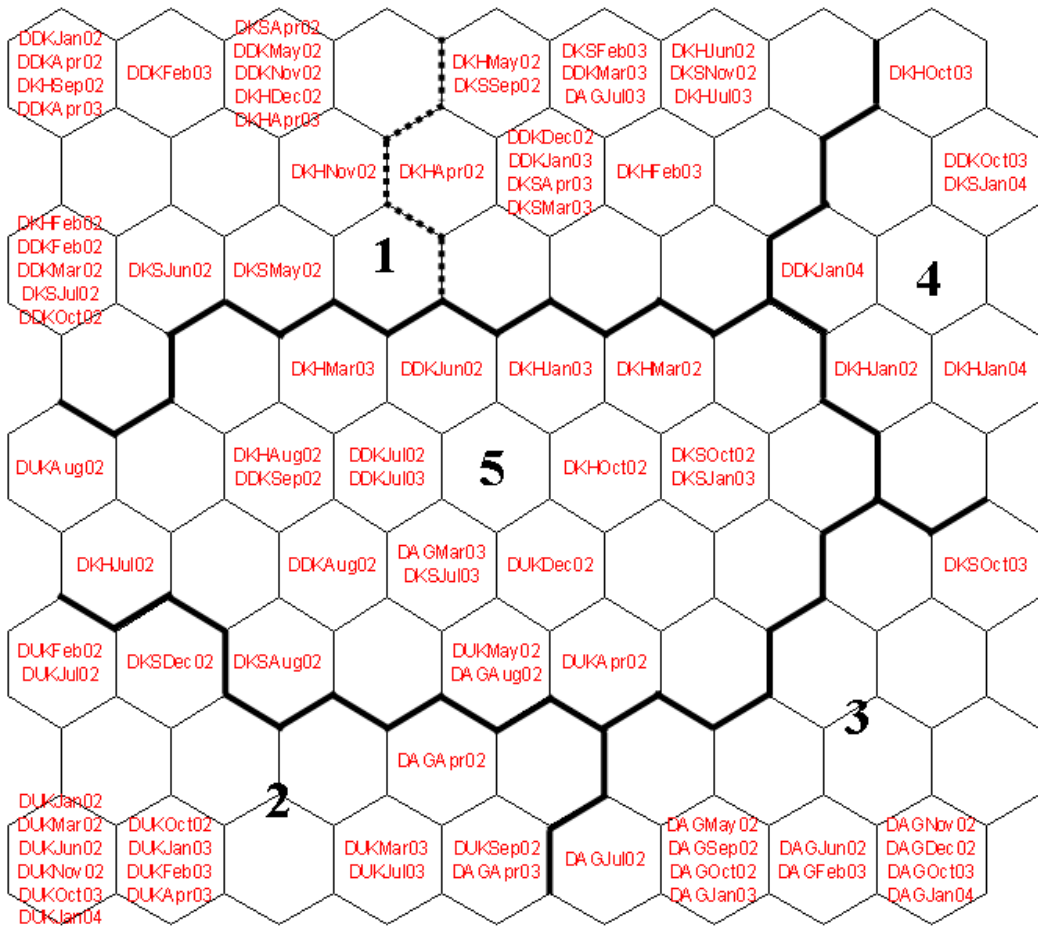


Fig. 4.8. SOM patterning of sampling sites based upon macroinvertebrate population density. Period: January, 2002 through January, 2004 (19 months); Sites: Duk, Ddk, Dks, Dkh, Dag (5 sites).

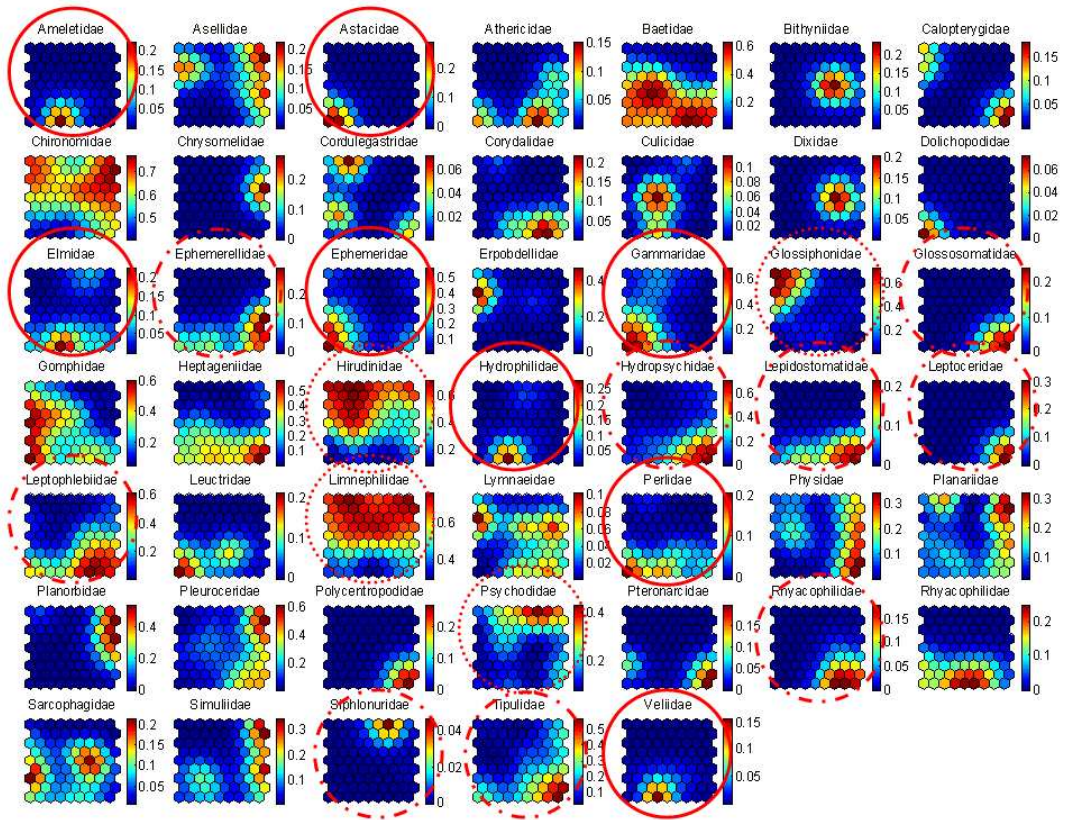


Fig. 4.9. Distribution of macroinvertebrate population on the SOM. Period: January, 2002 through January, 2004 (19 months); Sites: Duk, Ddk, Dks, Dkh, Dag (5 sites). Solid circle: reference site; Dotted circle: polluted site; Broken circle: Restoration sites.

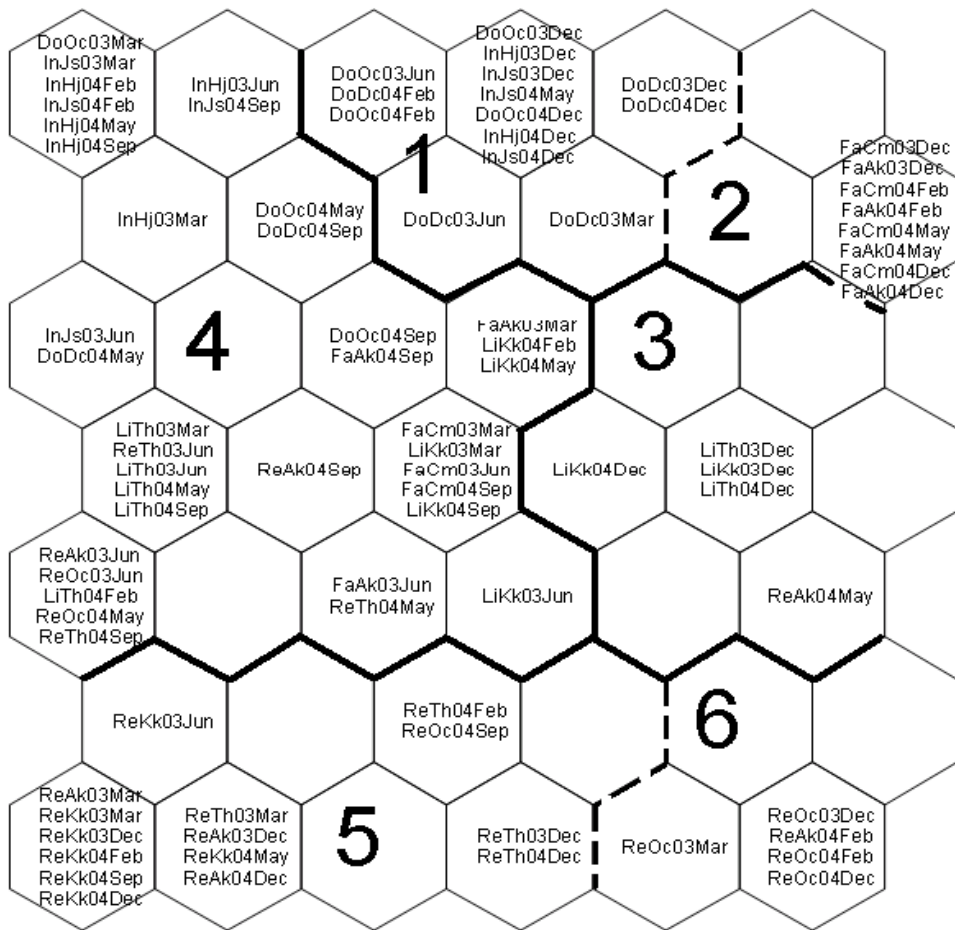


Fig. 4.10. SOM patterning of sampling sites based upon macroinvertebrate population density. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).



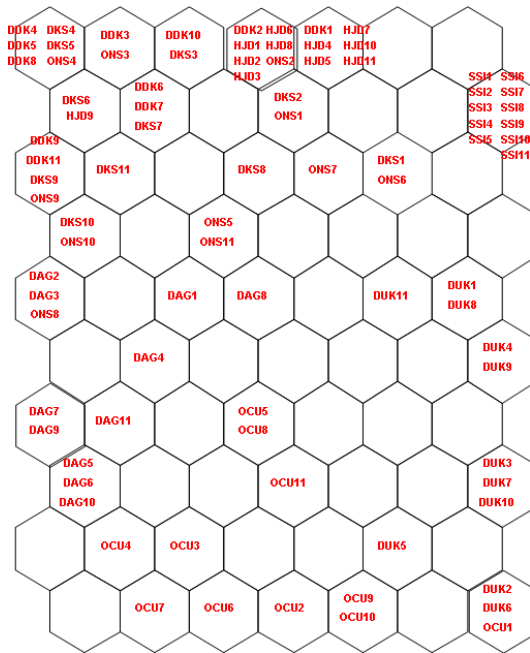


Fig. 4.12. SOM patterning of sampling sites based upon macroinvertebrate population density. Period: February, 2005 through April, 2007 (10 seasons); Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD (8 sites).

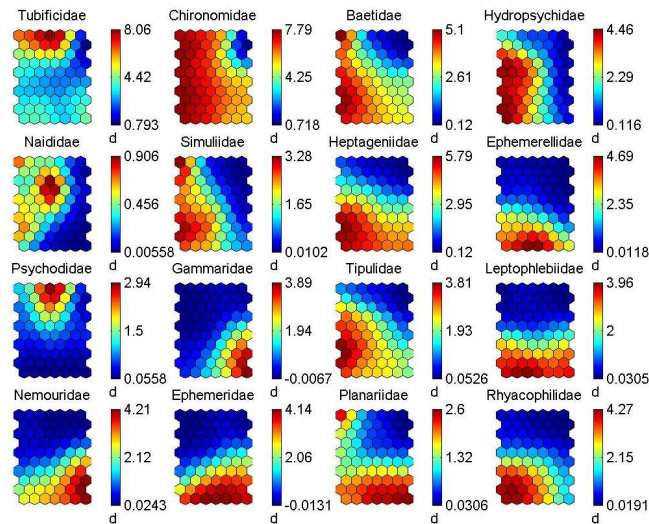


Fig. 4.13. Distribution of macroinvertebrate population on the SOM. Period: February, 2005 through April, 2007 (10 seasons); Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD (8 sites).

#### 4.3.1.3. 부착조류

부착조류 조사지점은 미생물 및 대형무척추동물 자료와 동일하나 조사시기는 다소 차이가 있다. 조사기간 동안의 대형무척추동물 출현량은 Fig. 4.14는 48개 지점(4개월 × 12개 지점)에서 출현한 44종의 부착조류 자료를, Fig. 4.16은 80개 56지점(7계절 × 8지점)에서 출현한 부착조류 8과의 자료를 이용하여 조사시기별 지점에 따른 그룹핑을 수행하였다.

Fig. 4.14에서는 전체적으로 5개의 클러스터로 나뉘어졌다. 전체적으로 오염의 종류에 따른 그룹핑 보다는 시기별로 그룹핑되는 경향이 강했으나 클러스터 4에서는 경작지와 축산폐수 오염지역이 일부 묶여져서 나타났다.

클러스터 1에서는 주로 *Staurastrum* sp., *Chroococcus* sp., *Oscillatoria* sp., *Synechocystis* sp.와 *Stephanodiscus* sp.가 우점하여 나타났다. 클러스터 2에서는 *Ankistrodesmus* sp., *Anabaena* sp., *Microcystis* sp., *Pseudanabaena* sp., *Aulacoseira* sp., *Cyclotella* sp., *Neidium* sp., *Pinnularia* sp., *Surirella* sp., *Cryptomonas* sp.와 *Euglena* sp.등이 주로 나타났다. 클러스터 3에서는 9종이 출현하였는데, 주로 와편모조류인 *Bacillaria* sp., *Ceminella* sp., *Gephyria* sp., *Stigeoclonium* sp. 등이었다. 클러스터 4에서는 6종이 우점하여 나타났는데, *Monoraphidium* sp., *Scenedesmus* sp., *Fragilaria* sp., *Melosira* sp., *Synedra* sp.와 *Tabellaria* sp.가 출현하였다. *Actinastrum* sp., *Cosmarium* sp., *Dictyosphaerium* sp., *Closterium* sp., 등은 클러스터 5에서 우점하여 나타났다.

*Acnmathes* sp., *Cymbella* sp. 와 *Trachelomonas* sp.는 클러스터 1과 2에서 모두 출현하였다. *Gomphonema* sp.는 클러스터 4와 5에서, *Nitzschia* sp.는 클러스터 2, 3, 5에서 높게 나타났다.

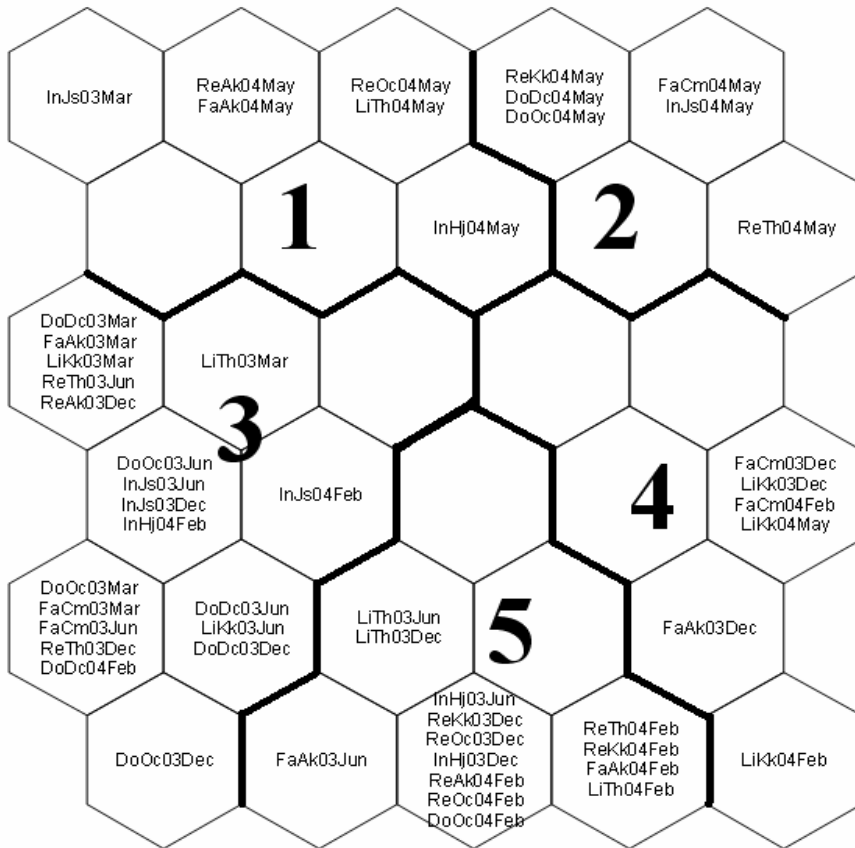


Fig. 4.14. SOM patterning of sampling sites based upon attached algal population density. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).



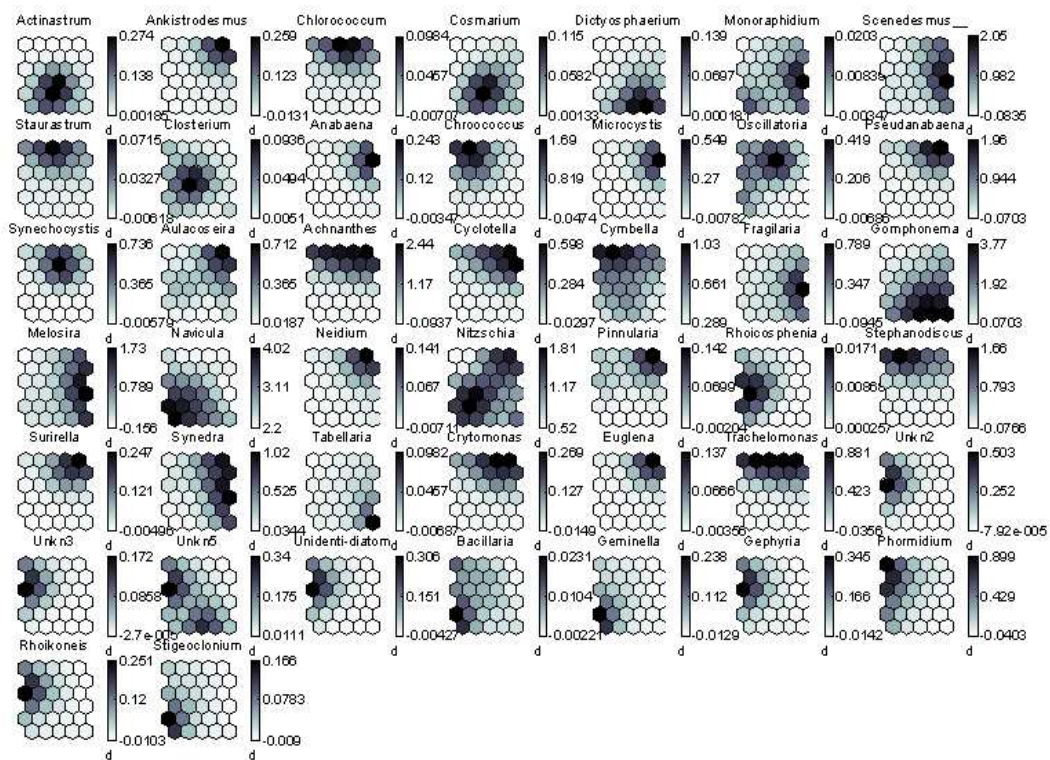


Fig. 4.15. Distribution of attached algal population on the SOM. Period: March, 2003 through December, 2004 (7 seasons); Sites: ReA, ReT, ReK, OCU, ONS, DUK, FaC, FaA, LiT, LiK, HJD, InJ (12 sites).

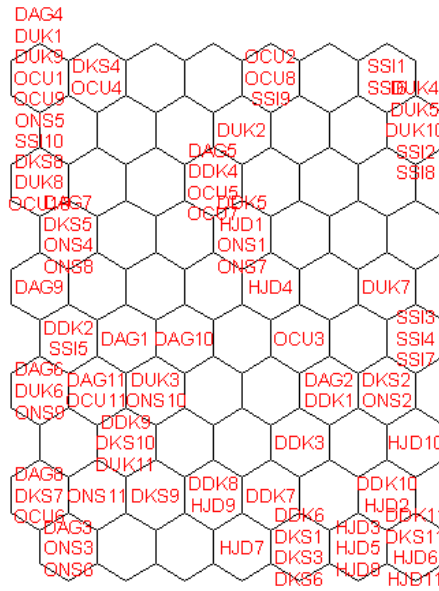


Fig. 4.16. SOM patterning of sampling sites based upon attached algal population density. Period: February, 2005 through April, 2007 (10 seasons); Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD (8 sites).

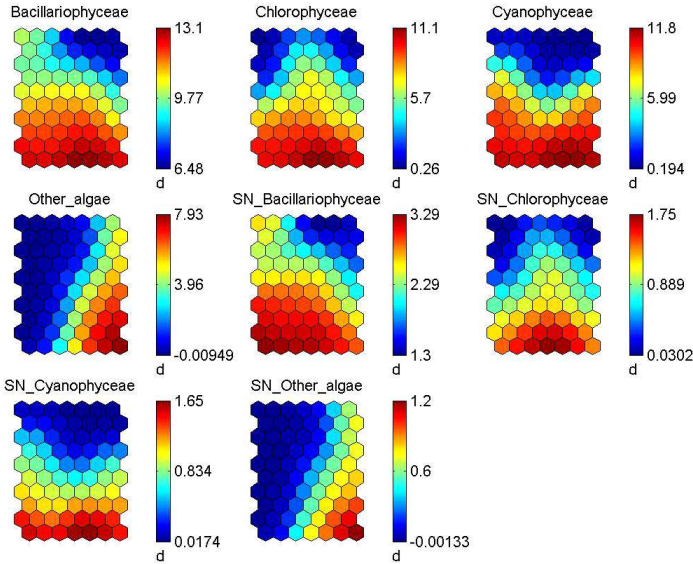


Fig. 4.17. Distribution of attached algal population on the SOM. Period: February, 2005 through April, 2007 (10 seasons); Sites: DUK, DDK, DKS, DAG, OCU, ONS, SSI, HJD (8 sites).

#### 4.3.2. 다층퍼셉트론을 이용한 생물군집과 환경요인의 연관성

다층퍼셉트론(Multi Layer Perceptron, MLP)을 이용하여 군집자료와 주요 환경요인과의 관련성을 검토하였다. Fig. 4.18은 실측 데이터와 예측된 데이터를 비교한 예인데 지수와 생물군집 및 환경요인과의 관련성을 나타낼 수 있었다. 다양한 자료에 대해 연관성을 구할 수 있었는데 본 자료의 경우 전체적으로 Pearson 상관계수가 0.77로 나타났다. 이를 통해 비선형적인 환경요인-생물 자료간의 관련성을 알 수 있는 모델 구성이 가능하였다.

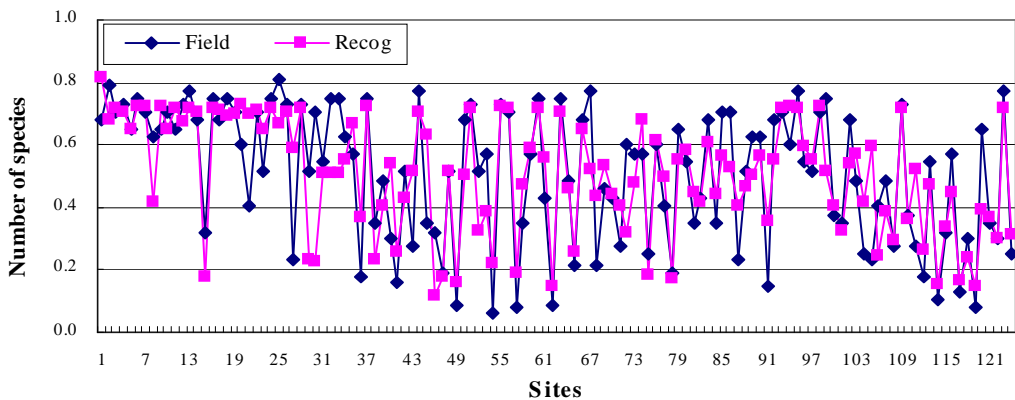


Fig. 4.18. Recognition of species richness using multi-layer perceptron.

다음 단계로 주요 환경자료를 이용하여 위 모델을 이용하여 주요 분류군에 대해 민감도 분석을 수행하였다. 종풍부도는 BOD가 증가함에 따라 감소하는 경향을 보였고, DO의 경우 반대 경향을 나타내었다(Fig. 4.19). 이러한 생물지수와 주요 환경자료와의 관련성은 조사된 생물자료가 전반적으로 이화학적 교란을 나타내는 환경지수와 일관성을 가짐을 나타내었다. 아울러 생태 조사를 통해 수집된 군집자료가 수질지수를 나타내는데 적절함을 시사하였다.

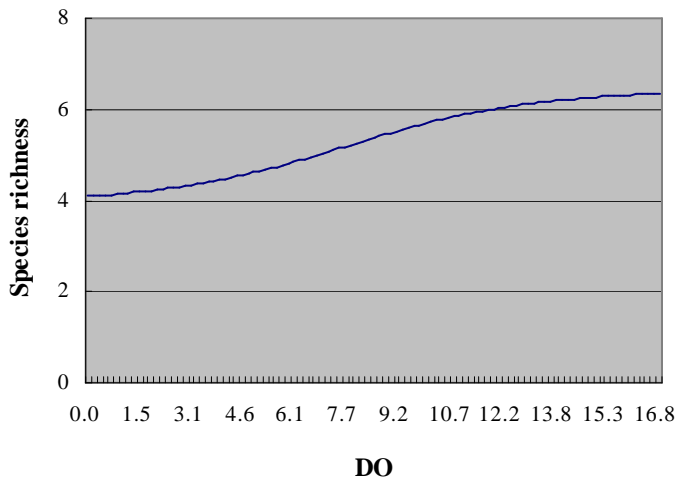
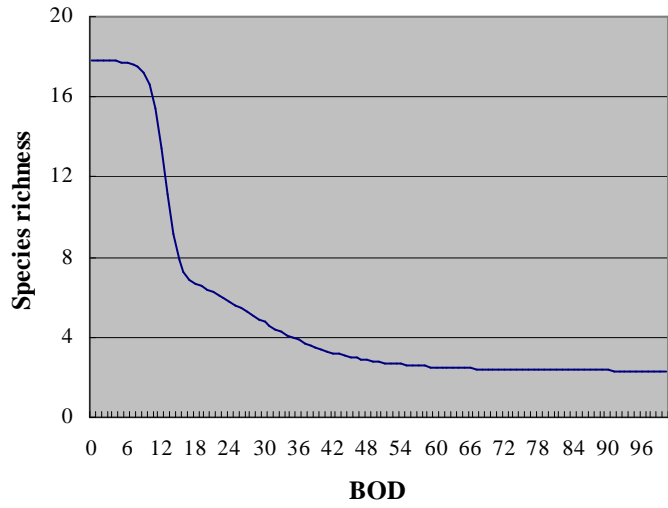


Fig. 4.19. Sensitivity analysis of species richness by MLP using BOD and DO as input data.

### 4.3.3. 다분류군 생물지수에 대한 환경요인과의 관계

환경요인에 대한 각 분류군의 반응을 보기위해, 학습된 다층퍼셉트론을 저서생물, 미생물 및 부착조류 군집에 적용하여 민감도를 분석하였다. 주요 환경요인인 COD와 BOD에 대한 각 분류군의 변이를 볼 수 있었다(Fig. 4.20, Fig. 4.21). 화학적 산소요구량인 COD에 대해서 부착조류 종풍부도(SR)가 큰 반응을 나타내었고, 반면 저서생물-SR은 큰 반응을 보이지 않았다.

생물학적 산소요구량인 BOD의 경우 세 분류군의 종풍부도에서 변이 폭이 커서, BOD의 변화에 따라 저서생물-미생물-부착조류의 중수가 민감하게 반응함을 알 수 있었는데, 상대적으로 미생물과 부착조류의 반응이 컸다. 이와같이 “군집지수”-“환경요인”간의 관련성이 분류군간에서 다양하게 나타났다. 이러한 다양한 관련성은 각 분류군이 환경영향을 각각 특성적으로 나타내었는데, 이를 통해 분류군은 독립적으로 생태계 양상을 표출함을 시사하였다.

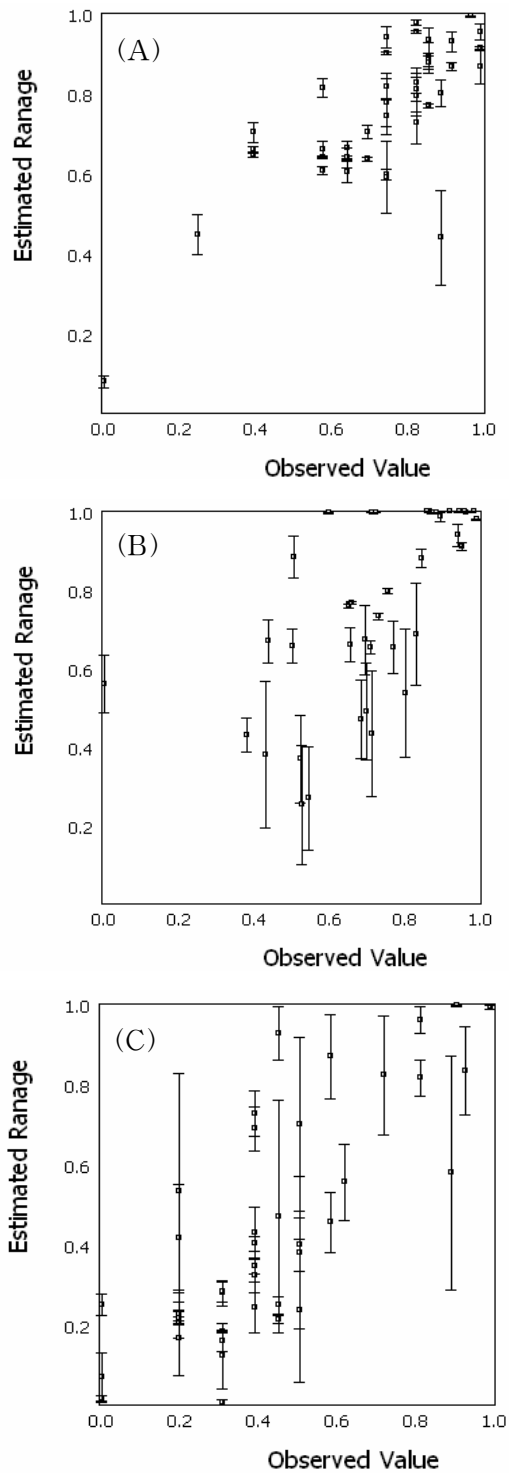


Fig. 4.20. Effect of COD on species richness of the inter-taxa. (A): Macroinvertebrate; (B): Microorganisms; (C)L Attached algae.

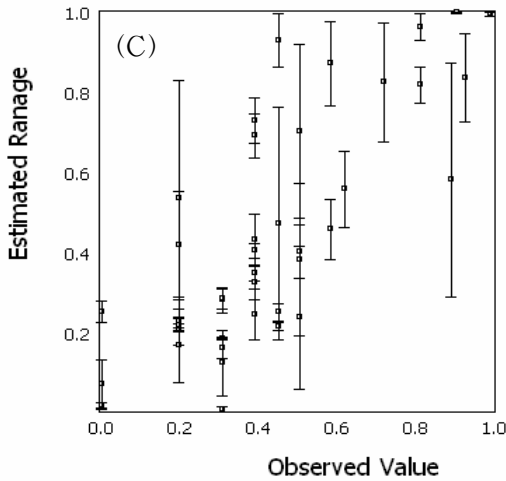
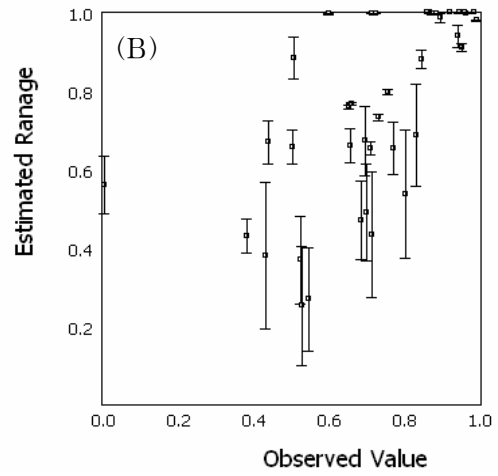
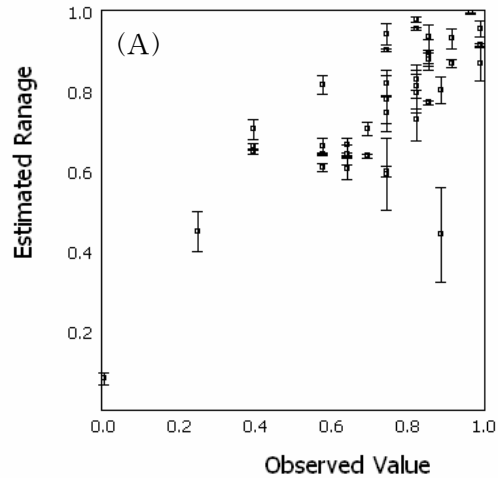


Fig. 4.21. Effect of BOD on species richness of the inter-taxa. (A): Macroinvertebrate; (B): Microorganisms; (C)L Attached algae.

#### 4.3.4. 각 분류군 생물지수의 연관성

생태자료를 통합하기 전에 각 분류군 생물지수의 연관성을 조사하였다. 각 분류군에서 이용되어진 지수는 아래와 같다. 사전에 생물지수 연구가 많이 수행되어진 저서성 대형무척추동물은 종풍부도(SRben), 총밀도(TDben), EPT 풍부도(EPTben), EPT 밀도(EPT-Dben), BMWP 지수(BMWpben), 다양도지수(Hben)를 검토하였다. 부착조류의 경우 종풍부도(SRalg)와 다양도지수(Halg) 및 총밀도(TDalg)를, 미생물의 경우 종풍부도(SRmic), 총밀도(TDmic)와 다양도지수(Hmic)를 적용하였다.

이들 지수들의 연관성을 보기위해 인공신경망인 SOM을 이용하여 유형화하였다(Fig. 4.22). 각 분류군의 주요 지수들은 별도의 그룹을 형성하여 독립적인 연관성을 나타내었다. 다만 TDalg는 부착조류군과 다소 떨어져 나타났다. 이는 각 분류군이 독립적으로 수질을 나타냄을 시사하였다.

분류군 지수는 아래의 Table 4.5에서 보여지는 각 생물지수간의 Pearson 상관관계수에서도 잘 나타났다. 상관관계수 검토에서 저서성 대형무척추동물과 조류의 밀도(TDben, TDalg)는 다른 지수와 관련성이 낮게 나타나 대상에서 제외하였다. 분류군내에서는 전체적으로 상관관계수가 높았으나, 다른 분류군과의 지수와는 낮게 나타났다. 저서생물내의 지수가 분류군내에서 일관적으로 높게 나타났다. 대부분 0.7이상으로 나타났으며 BMWP, EPT가 전반적으로 높게 나타났다. 그러나 미생물과 조류에서 구해진 지수와의 상관관계수는 낮았다. 조류에서도 조류 내의 상관관계수는 상대적으로 높게 나타났다. 그러나 전반적으로 저서생물 지수에서 보여진 상관관계수보다는 낮았다. 미생물의 경우는 분류군내의 지수는 높게 나타났다. 특히 Hmic와 SRmic는 높게 나타났다. 조류와 미생물에서도 분류군간 상관관계수는 매우 낮았다.

Fig. 4.22와 Table 4.1에 의하면 앞서 언급하였듯이 각 지수들은 분류군에 따라 다르게 나타남을 알 수 있었다. 이는 오수생태계의 주요분류군의 지수는 별도로 수질을 나타냄을 시사하며, 이는 조류, 저서생물, 미생물이 각각 생산자, 소비자, 분해자임을 고려할 때 독립적으로 수질과 연관성을 가짐을 나타내었고, 통합적 수질지수를 나타내기 위해서는 각 분류군을 별도로 포함시키는 것이 의미가 있음을 시사하였다.



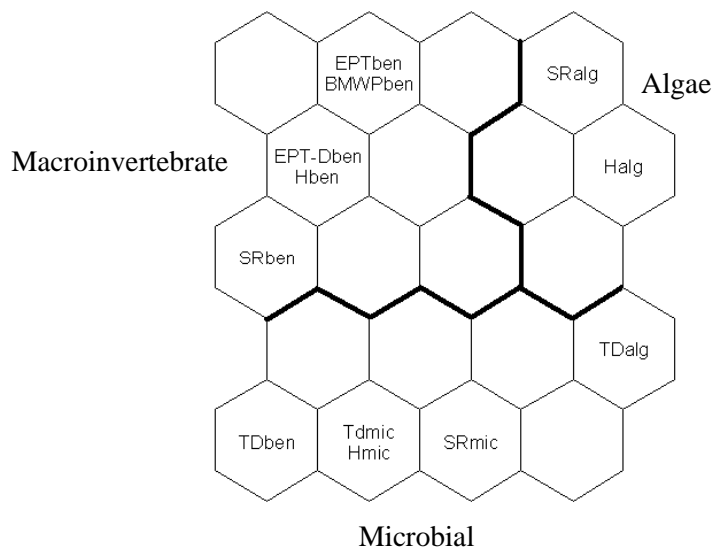


Fig. 4.22. Grouping of indices for the inter-taxa by SOM.

Table 4.1. Pearson correlation coefficients for indices of the inter-taxa.

|          | <i>Macroinvertebrates</i> |         |        |        |        | <i>Algae</i> |         | <i>Microorganisms</i> |        |
|----------|---------------------------|---------|--------|--------|--------|--------------|---------|-----------------------|--------|
|          | SR                        | EPT     | EPT-D  | BMWP   | H      | SR           | H       | SR                    | TD     |
| SRben    |                           |         |        |        |        |              |         |                       |        |
| EPTben   | 0.814*                    |         |        |        |        |              |         |                       |        |
| EPT-Dben | 0.772*                    | 0.904*  |        |        |        |              |         |                       |        |
| BMWPben  | 0.875*                    | 0.910*  | 0.867* |        |        |              |         |                       |        |
| Hben     | 0.658*                    | 0.832*  | 0.700* | 0.781* |        |              |         |                       |        |
| SRalg    | -0.050                    | 0.012   | 0.023  | -0.051 | -0.065 |              |         |                       |        |
| Halg     | 0.154                     | 0.169   | 0.231  | 0.131  | 0.059  | 0.730*       |         |                       |        |
| SRmic    | -0.333*                   | -0.263* | -0.230 | -0.240 | -0.092 | -0.200       | -0.329* |                       |        |
| Tdmic    | -0.209                    | -0.119  | 0.007  | -0.050 | 0.001  | -0.341*      | -0.328* | 0.727*                |        |
| Hmic     | -0.288*                   | -0.250  | -0.228 | -0.223 | -0.069 | -0.193       | -0.314* | 0.971*                | 0.687* |

\* Values indicate the case of  $p < 0.05$ .

각 분류군 생물지수와 환경자료와의 상관계수를 Table 4.2에 나타내었다. 이 경우, 부착조류의 분류가 집중적으로 이루어진 2004년 9월과 12월 자료를 이용하였다. 저서성 대형무척추동물 지수의 경우 다른 조류와 미생물 지수에 비해 환경자료와의 연관성이 높았는데, 특히 BMWP와 다양도지수(Hben)가 BOD, TOC, TP, TN 등의 환경요인과 상관계수가 일관적으로 높게 나타났다. 조류자료를 이용해 구해진 DA<sub>Ipo</sub> 지수도 BOD, TOC 등의 환경요인과 연관성이 높게 나타났다. 그러나 본 연구의 여건상 부착조류의 분류에 어려움이 있고 시간이 소요되어, 통합수질지표에는 DA<sub>Ipo</sub> 지수를 응용한 지수를 적용하였다.

Table 4.2. Pearson correlation coefficients for environmental parameters and biological indices.

| <i>Biotic indices</i>  | <i>Environmental factor</i> |        |       |        |                                 |        |        |
|------------------------|-----------------------------|--------|-------|--------|---------------------------------|--------|--------|
|                        | BOD                         | TOC    | COD   | T-P    | NO <sub>3</sub> <sup>-</sup> -N | TKN    | TN     |
| SR-ben                 | -0.29                       | -0.47* | -0.42 | -0.45* | -0.43                           | -0.46* | -0.53* |
| EPT-ben                | -0.37                       | -0.42  | -0.05 | -0.47* | -0.33                           | -0.41  | -0.57* |
| BMWP-ben               | -0.39                       | -0.47* | -0.33 | -0.47* | -0.50*                          | -0.46* | -0.58* |
| H-ben                  | -0.52*                      | -0.49* | -0.01 | -0.65* | -0.35                           | -0.48* | -0.50* |
| SR-micro               | -0.33                       | -0.14  | 0.38  | -0.25  | -0.21                           | -0.14  | -0.09  |
| Density-micro          | -0.18                       | -0.36  | 0.00  | 0.28   | -0.37                           | -0.36  | -0.37  |
| SR-alg                 | -0.19                       | 0.15   | -0.09 | -0.10  | -0.06                           | 0.15   | 0.31   |
| Density-alg            | 0.03                        | 0.39   | 0.17  | 0.32   | 0.27                            | 0.38   | 0.46*  |
| H-alg                  | -0.24                       | -0.19  | -0.03 | -0.30  | -0.04                           | -0.19  | 0.10   |
| DA <sub>Ipo</sub> -alg | -0.50*                      | -0.53* | -0.29 | -0.67* | -0.47*                          | -0.52* | -0.43  |

\* Values indicate the case of  $p < 0.05$ .

#### 4.3.5. 다분류군의 군집지수 연관성

각 분류군 자료를 통합적으로 유형화하기 위해 주요지수를 상호 비교하였다. 주요 군집지수인 다양도(Hellawell, 1986)를 삼각형 그래프에서 비교하였다(Fig. 4.23). 조사된 지수는 각 분류군 내에서의 최대치를 기준으로 0.0 - 1.0 사이에서 정규화 하였다. 전체적으로 지수간 변이 폭을 알 수 있었는데, Fig. 4.23에서 볼 수 있듯이 다양도 지수는 다른 분류군에 비해 저서성 대형무척추동물과 병행하여 넓게 분산되어 있어, 저서성 대형무척추동물이 생태계 변이 폭을 크게 나타냄을 알 수 있었다. 다양도 지수는 저서성 대형무척추동물에서 0.1 - 0.7 사이에서 변이를 나타내었고 조류와 미생물은 각각 0.3 - 0.8, 0.4 - 0.8로 약간 높은 수준에서 상대적으로 좁게 나타났다.

각 조사지에 따라 삼각형 그래프(Fig. 4.23)에서 각 지수의 연관성은 다양하게 나타났다. 대조조사지에서는 저서생물의 다양도가 높았고, 미생물과 부착조류의 다양도는 중간정도의 수치를 나타내었다. 오염이 심하게 된 경우에는(유기물과 산업폐수 오염지역) 저서생물의 다양도는 낮아지고, 미생물과 부착조류의 다양도가 상대적으로 높게 나타났다. 오염정도가 상대적으로 낮은 경작지와 축산폐수 오염지역은 저서생물의 다양도는 높아지고 다른 분류군 다양도는 상대적으로 낮아지는 중간단계 유형을 보였다. 이로써, 각 분류군의 다양도 지수의 조합은 오염정도에 따라 다양하게 반응함을 알 수 있었다.

각 분류군의 종풍부도에 대해서도 위와 유사한 방법으로 분석하였다(Fig. 4.24). 전반적으로 다양도 경우와 유사했으나 저서성 대형무척추동물과 조류-미생물간의 종풍부도 범위차는 다소 좁혀졌다. 대조조사지에서 저서생물과 미생물의 종풍부도가 높게 나타났고, 부착조류의 경우 낮게 나타났다. 오염이 심한 경우에는 저서생물의 종풍부도는 낮아지고, 부착조류는 증가하였다. 그러나 미생물의 종풍부도는 오염에 관계없이 비교적 일정하게 나타났다.

이와 같이 주요 군집지수를 생산자, 소비자, 분해자의 관점에서 종합적으로 유형화 할 수 있었다. 위에서 언급한 바와 같이 각 조사지의 교란 양상은 세 분류군간의 관련성에서 각기 특성적인 위치를 보임을 알 수 있었다. 이로써 수질변이를 나타냄에 있어 생산자, 소비자, 분해자를 대표하는 각 분류군을 독립적으로 검토할 필요가 있음을 알 수 있었다.

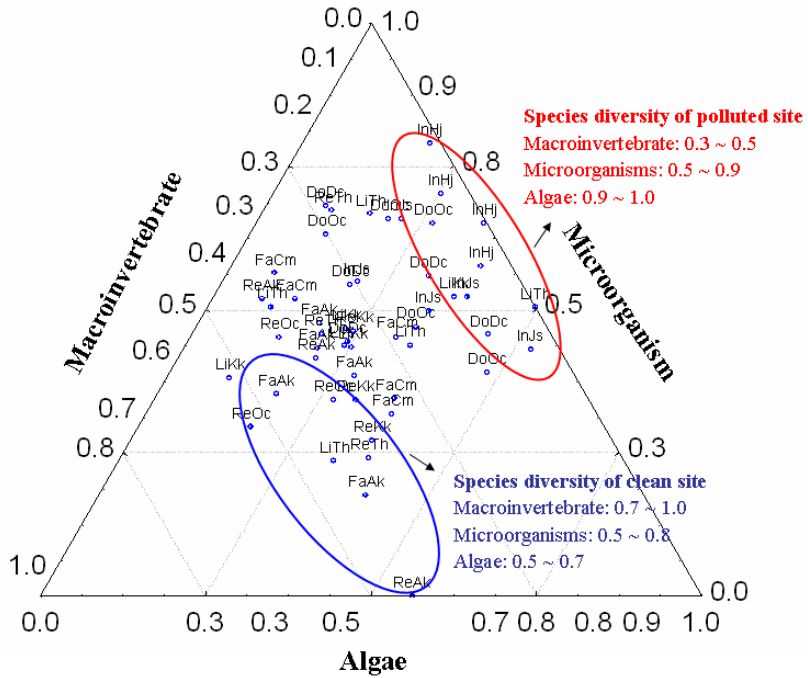


Fig. 4.23. Distribution of each taxon depending on species diversity.

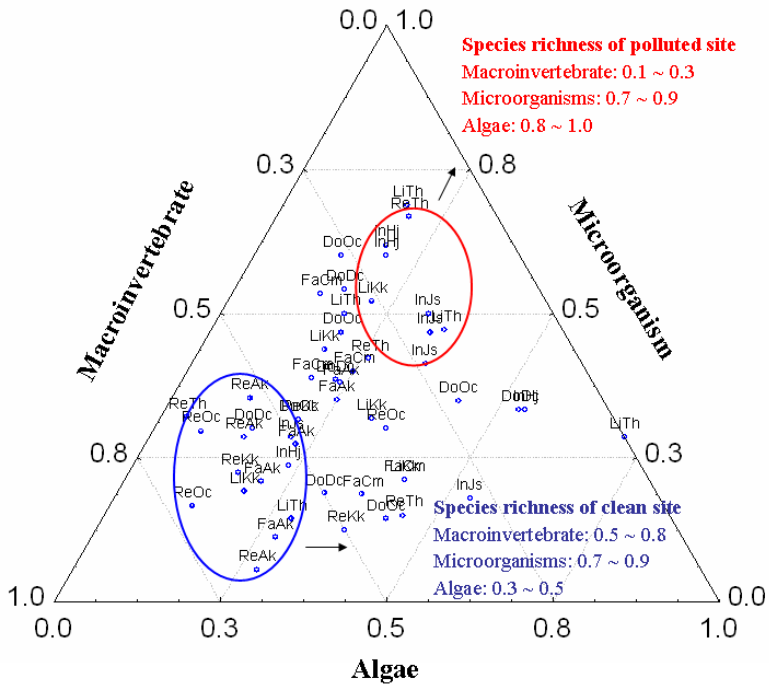


Fig. 4.24. Distribution of each taxon depending on species richness.

#### 4.4. 결론

본 연구에서는 통합수질지수 개발을 위해 각 오염종류에 따른 분류군의 군집을 유형화하고 민감도를 분석하여 환경요인과 다분류군 군집구조간의 상관관계를 파악하고자 하였다. 그 결과 통합수질지수를 개발하기 위해서는 분류군을 대표할 수 있는 지수 선정이 중요하며, 각 분류군의 독립적인 연관성을 고려하여함을 확인하였다.

1. SOM 분석결과 대형무척추동물의 경우 오염지별 군집 양상이 뚜렷이 구분되는 것으로 나타났으나 조류 및 미생물 군집의 경우 지역이나 오염원(오염도) 보다 계절적 영향에 더 민감한 것으로 나타나서 분류군간의 특성을 인정해야 할 필요가 있는 것으로 보인다.
2. 환경자료와 각 분류군 간의 상관성 분석결과 세 분류군 모두가 환경요인(COD 및 BOD)에 대해 민감하게 반응하였으나, 상대적으로 미생물과 조류의 반응이 컸다. 이는 각 분류군이 환경영향을 각각 특성적으로 나타냄으로써 각 분류군은 독립적으로 생태계 양상을 표출함을 시사한다.
3. 각 분류군을 대표하는 지수들을 이용하여 SOM 분석한 결과 지수들은 분류군별로 별도의 그룹을 형성하여 독립적인 연관성을 나타내었다. 이를 통해 각 분류군이 독립적으로 수질을 나타냄을 재확인하였다.
4. 각 분류군을 대표하는 지수들 간의 상관성을 분석한 결과 대형무척추동물과 조류 지수들 간에는 상관성이 있는 것으로 나타났으나 기존 미생물 지수와의 상관성은 확인할 수 없었다. 이는 분석에 이용된 미생물지수가 환경을 반영하지 못하데 기인하는 것으로 사료된다.
5. 통합수질지수를 개발하기 위해서는 각 분류군이 독립적으로 수질을 반영하는 특성을 고려할 필요가 있으며, 각 분류군과 환경과의 상관성을 규명하고 각 분류군의 특성을 가장 잘 반영할 수 있는 지수의 개발이 필수적이다.

## 제 5장 통합지수 개발을 위한 미생물 지수 개발

본 연구에서는 임의의 미생물 밀도가 특정 환경적 요인과 밀접한 관련이 있을 것으로 가정하고 관측된 환경요인과 각각의 미생물들의 상관성을 조사하였다. 환경요인과 미생물들 간의 상관성이 높을 경우 회귀분석을 통해 이들 간의 선형관계식을 얻을 수 있고, 이 선형관계식은 주어진 환경요인에 대한 특정 미생물의 밀도를 나타낸 것이라 볼 수 있으므로, 각 미생물 개체들에 대한 회귀식을 비교 분석하여 미생물들을 지수화하고자 하였다.

그러나 미생물 군집구조와 환경요인간의 상관성 분석이 어려웠다. 이에 본 연구에서는 주어진 환경조건에서 출현하는 미생물의 종류 및 출현패턴을 확률 모델에 기반하여 분석한 후 그 결과로부터 환경 상태를 잘 반영하는 미생물을 선별하고 이를 특성화하여 미생물 지수화를 시도하였다. 데이터 분석절차는 다음과 같다. 먼저 요인분석을 통해 오염도와 관련된 환경변수들을 선별하고 이들의 공통요인을 추출하여 이들로부터 요인점수(Factor scores)를 산출하였다. 요인분석에 의해 선별된 환경변수들을 요인점수의 축 선상에서 비교 분석하여 관측 지점별로 순위를 지정하고, 요인점수의 구간별 밀도에 기초하여 두 개의 그룹으로 분류하였다. 즉, 두 개의 그룹은 요인점수가 높은 오염지역 및 요인점수가 낮은 청정지역을 의미한다. 각 그룹별로 요인점수에 대한 개개의 미생물의 상관성 또는 미생물 밀도를 조사하여 청정지역에서 관측되는 미생물, 오염지역에서 관측되는 미생물, 청정지역과 오염지역에서 모두 관측되는 미생물로 분류하고 각각의 주요도 순위를 지정한 후, 지정된 순위에 따라 미생물의 지수화를 수행하였다.

## 5.1. 연구배경

수질평가를 위한 생물평가 연구는 주로 군집연구에 국한되어 있다. 국내외에서 어류, 수서곤충, 대형무척추동물 및 조류에 대한 연구가 활발히 이루지고 있어 수질평가를 위한 다양한 지표 및 지수들이 제시되고 있고, 어류를 이용한 IBI, 부착조류를 이용한 DAIPo 및 TDI 그리고 대형무척추동물을 이용한 BMWP, TBI, 및 BS 지수들이 대표적인 예이다. 그러나 타 분류군에 비해 민감도 및 특이성에서 우수한 미생물을 이용한 수질평가 연구는 거의 전무한 실정이다. 또한 현재까지 수질평가는 이화학적 지표나 생물지수 중 중요 분류군에 대한 별도의 연구가 진행되어 다분류군을 이용한 통합적 수질지표개발 연구는 초기단계이다. 이에 다분류군(미생물, 조류 및 대형무척추동물)을 이용한 통합수질지수를 개발을 통해 생태계의 오염도는 물론 생태계의 변화과정 추적에도 활용할 수 있는 기반이 마련되어야 한다. 통합수질지수를 개발하기 위해서는 다양하고 복잡한 수질 정보에 대한 합리적이고 신뢰성 있는 해석을 위해서는 수질자료가 가지는 오차를 줄이고 정보의 손실을 최소로 하는 과학적인 해석방법이 필요하다.

최근 다변량분석법 및 인공신경망을 이용 수질평가 연구가 활발히 진행되고 있다. 다변량 분석법은 요인분석, 주성분분석 및 군집분석 등과 같은 여러 방법이 있다. 국내에서는 다변량분석법을 이용한 수질평가를 통해 수질에 영향을 미치는 인자들을 분석하는 연구들이 다수 수행되고 있다(신 등, 1998; 김, 2002; 유, 2002; 김, 2003; 김 등, 2007). 주성분분석 및 요인분석을 통해 호소의 영양상태, 수질 관련 항목들의 평가와 관리를 위한 연구가 국내외적으로 수행되었다(Petersen, et al., 2001; Bernare, et al., 2004). 인공신경망을 적용한 생태자료 연구는 1990년대부터어 생물의 생태 해석을 위해 인공신경지능망인 Kohonen network를 이용하여 생태군집을 분류하고(Chon, et al., 1996), 나아가 군집의 시간적 변이에 대한 유형화, 분류군간 상관성 분석 연구가 시도된 바 있다(Chon, et al, 2000; Kim, et al., 2008). Rockhow(1999)는 확률신경망을 적용한 수질 예측 연구를 수행한 바 있으며, Dimictic Glacial/Boreal Lakes 수질관리를 위해 확률밀도함수 분석을 통하여 수질지표들(pH, alkalinity, conductivity, hardness, color, Secchi depth, and total phosphorus)의 특성화 연구가 수행된 바 있다(Thierfelder, 2005).

## 5.2. 실험방법

### 5.2.1. 데이터 분석전략

미생물 지수화를 위한 데이터는 크게 환경평가를 위한 환경데이터와 각각의 환경조건(오염도)에 따른 미생물 데이터로 대별된다. 환경평가를 위해 수집된 데이터는 13개 항목의 이화학적 환경변수(BOD, T-N, T-P 및 TOC 등)에 대하여 시료채취 시기 및 지역에 따라 각각 253개의 수치 값으로 기록되어 있다. 미생물 데이터는 PCR-DGGE 분석 기법에 의해 추출된 115종의 미생물에 대한 출현밀도에 대응되는 수치 값이 기록되어 있다. 환경 데이터는 수집여건에 따라 누락된 경우가 있으나, 상관성 분석 등의 기본적인 통계적 기법을 적용하여 데이터 전반에 걸친 추이를 관찰하는 것을 위한 목적으로는 크게 무리가 없었다. 반면, 미생물 데이터는 미생물의 종류가 매우 다양하며 조류 및 저서생물과는 달리 특정 환경에 종속적으로 출현되는 경향이 있었으며, 특히 PCR-DGGE 분석에 드는 비용 및 시간적 한계로 인해 상대적으로 누락된 부분이 많기 때문에 직접적인 상관관계를 규명하는 것은 쉽지 않았다. 다시 말해, 수집된 환경-미생물 데이터는 종합적으로 고려해 볼 때 불완전하며 샘플 수에 비해 차원이 매우 크다는 문제가 있었다. 이러한 특성의 데이터를 회귀분석이나 상관성분석과 같은 기존의 통계학적 방법으로 분석하는 것은 매우 까다로울 뿐 아니라 분석결과에 있어서도 높은 신뢰성을 기대하기 어렵다. 또한 본 연구에서 다루고자 하는 미생물 지수화는 주어진 환경-미생물 데이터의 분석결과로부터 결정된다고 간주할 수 있으므로 효과적인 분석방법의 선정은 매우 중요하다.

이에 환경 및 미생물 데이터의 분석 및 분석결과로부터 미생물 지수화 및 환경평가를 효과적으로 수행하기 위해 다음과 같은 절차에 따라 분석을 수행하였다. 우선 13개의 차원을 가지는 환경데이터로부터 가능한 두 개 이하의 차원을 가지도록 각 환경변수를 대표할 수 있는 공통성분을 요인분석에 의해 추출하였다. 이때 공통성분은 환경상태(오염도)를 잘 설명하며, 공통성분과 환경 변수 간에는 회귀분석에 의한 관계식이 주어진다. 분석의 용이성을 위해 추출된 공통요인에 할당된 점수는 오염정도를 반영한다고 가정하고 미생물 데이터의 분석 시 각각의 환경변수와 미생물 변수의 관계를 비교분석하는 대신 공통요인만을 참조데이터로 활용하였다. 다음으로 미생물과 공통요인간의 관계를 규명하



기 위해 공통요인에 대한 데이터를 두 개의 그룹으로 분류하고 각각의 그룹에 대한 미생물의 출현밀도를 확률모델로 근사한 후 이를 비교분석하여 미생물을 특성화함으로써 그 출현패턴을 분석하였다. 이를 통해 미생물은 각각의 특성과 출현패턴에 대한 정보를 지니게 되며 이는 환경오염정도에 따른 미생물 출현에 대한 확률모델로부터 얻어진 결과이기 때문에 미생물의 지수화에 활용될 수 있다.

미생물에 지수를 부여하는 과정에서 상대적으로 우수한 오염도 예측성능 및 환경평가의 영향력 등이 잘 반영될 수 있도록 각각의 미생물 군집을 대상으로 상대적인 평가를 실시하였다. 또한 최종적으로 지수가 부여되는 미생물은 일련의 단계를 거쳐 자동적으로 선택되고 적합한 지수점수가 부여될 수 있도록 하여 결과의 객관성을 높일 수 있도록 하였다.

### 5.2.2. 요인분석에 의한 환경요인의 추출 및 환경변수와의 관계

환경데이터는 관측된 샘플 수(253 sites)에 비해 상대적으로 차원이 높기 때문에 분석결과의 신뢰성을 높이기 위해서는 영향력이 적은 변수를 제거하여 필요한 변수들만을 선택한다거나(variable selection) 변수 간의 상관성 등에 기반한 차원축소(dimensionality reduction)기법 등이 고려되어야 한다(Wolf and Bileschi, 2005). 요인분석은 여러 개의 변수들이 서로 어떻게 연결되어있는 가를 분석하여 이들 변수간의 공통요인(common underlying dimensions)을 이용하여 설명하는 다변량분석 기법이다(이영준, 2002). 따라서 고차원의 환경변수들에 대해 요인분석을 시행하면 각 환경변수들의 특징을 가장 잘 설명할 수 있는 저 차원의 숨은 요인들(Latent Factors)을 추출가능하며 이러한 개념을 도시화 하면 Fig. 5.1과 같다.

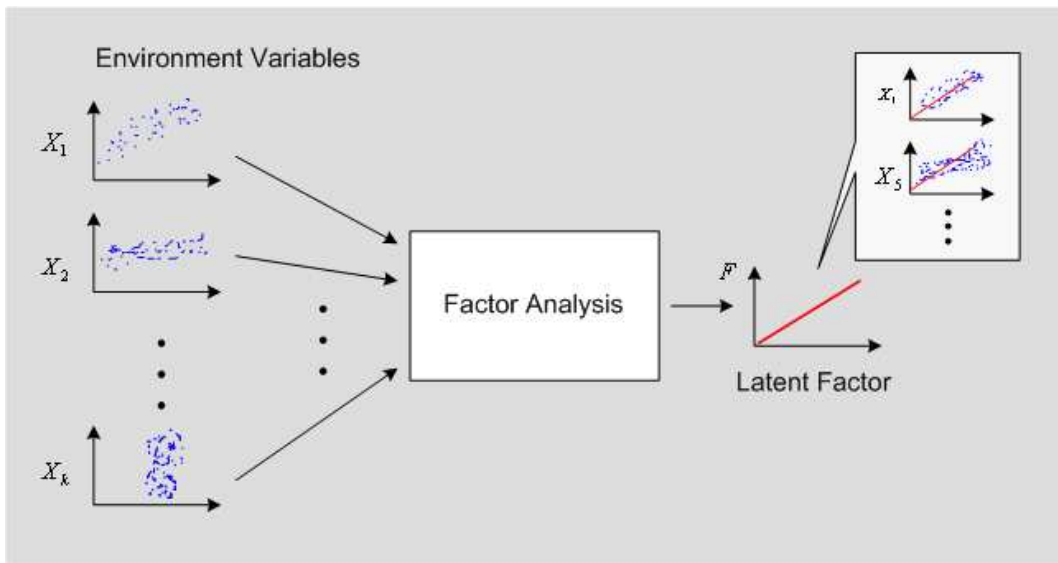


Fig. 5.1. Dimensionality reduction and principal-factor extraction through factor analysis technique.

일반적으로 요인분석은 다음과 같은 요인모형을 가정한다.  $k$ 개의 정규화된 변수  $z_1, z_2, \dots, z_k$  가 변수  $x_1, x_2, \dots, x_k$ 로부터 주어질 때 요인모형은 숨은 요인 (Latent factors)  $F_1, F_2, \dots, F_m$ , ( $m \leq k$ )들의 선형결합으로 표현 가능하다.

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{k1} & w_{k2} & \cdots & w_{km} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_k \end{pmatrix} \quad (5.1)$$

여기서  $w_{ij}$  는  $i$ 번째 변수에 대한  $j$ 번째 요인의 요인적재량(Factor loading)이라고 하며  $F_j$ 는  $j$ 번째 공통요인(Common factor)을 나타낸다.

본 연구에서는 SPSS를 이용하여 주어진 13개의 환경 변수 전체를 대상으로 요인분석을 수행하였고 요인을 추출하기 위한 방법으로 주성분 분석법(PCA: Principal Component Analysis)을 이용하였다. 본 연구에서 요인분석의 목적이 미생물 데이터를 분석하기 위한 오염정도를 잘 설명할 수 있는 참조자료를 얻기 위한 것이므로 다음과 같은 절차를 통해 최종적인 변수 및 요인점수를 추출하였다. 먼저 모든 환경변수를 대상으로 요인분석을 실시하여 주요 요인들을 추출하였다. 이때 모든 환경변수는 실측값을 그대로 사용하지 않고 로그공간으로 변환한 후 이를 평균이 0이고 분산이 1이 되도록 정규화 하여 사용하였다.

일반적으로 요인분석을 통해 얻은 요인들의 차수는 주요요인들의 고유치를 기준으로 정해지는데 각 요인의 고유치(Eigen value)는 요인을 설명하는 분산의 크기를 나타내는데 이 고유치가 1이상일 경우에 하나의 요인이 변수 1개 이상의 분산을 설명한다. 반면, 고유치가 1이하인 경우 1개의 요인이 변수 1개의 분산을 설명할 수 없으므로 고유치가 1.0 이상이 되는 요인만을 추출하면 하나 이상의 변수를 설명할 수 있는 의미 있는 요인들을 얻을 수 있다. 이 과정에서 제거된 요인에 대해서는 상관성이 높고 선택된 요인에 대해서는 상관성이 낮은 변수가 존재하면 이 변수는 다음 요인분석에서 제외한 후 나머지 선별된 변수들에 한해 요인분석을 다시 수행하였다. 이때 추출된 요인이 되도록 많은 변수와 공통성을 보이며 이들 변수가 환경오염과 밀접한 관련이 있는지 확인하는 과정을 반복수행하여 주요 요인들을 추출하였다.

상기의 과정을 통해 추출된 요인은 선정된 환경변수들을 대표하는 것으로 판명되었고 이들의 선형관계식을 통해 요인차원으로 사상될 수 있었다. 또한 요

인과 각 환경변수들은 독립적인 회귀분석을 통해서 관계식을 얻을 수 있으므로 요인점수만으로 선정된 환경변수들의 값을 추정할 수 있으며 그 반대의 경우도 가능하다.

### 5.2.3. 환경요인에 대한 확률모델 및 오염지역 분류

요인분석으로부터 추출된 공통요인 점수는 미생물의 출현패턴을 분석하기 위한 자료로 활용되었다. 그러나 5.2.1절에서 설명한 바와 같이 공통요인과 미생물간의 상관관계를 직접적으로 규명하는 것은 쉽지 않았다. 따라서 본 논문에서는 요인-미생물의 관계를 확률모델을 통해 근사화한 후 확률적 추론을 통한 분류문제로서 접근하여 이들의 관계를 분석하고자 하였다.

요인분석을 통해 추출된 요인이 수질의 상태를 잘 반영한다고 가정하면 요인 점수의 분포로부터 몇 개의 클래스를 분류한 후 분류된 클래스에 해당되는 미생물의 출현빈도수를 조사하면 수질상태에 대한 미생물의 출현패턴을 어느 정도 파악이 가능하다. 본 연구에서는 각 클래스에 해당되는 미생물의 출현패턴을 확률모델을 통해 분석하였고 각 패턴을 분류하는데 있어 확률적 추론방법을 적용하여 상대적으로 결측값(Missing value)이 많은 미생물 데이터를 효과적으로 분석하고자 하였다. 이러한 경우와 같이 관측된 정보에 결측값이나 노이즈가 포함된 정보가 주어질 경우 확률이론과 결정이론을 조합하여 주어진 정보에 대한 최적의 예측치를(Optimal predictions) 구할 수 있다(Bishop, 2006). 이에 본 연구에서는 베이저안 이론을 기반으로 요인 및 미생물 데이터의 확률적 모델 및 추론을 통해 미생물의 분포패턴을 분류하여 미생물의 지수화를 위해 필요한 정보를 얻고자 하였다.

$d$ 차원-랜덤변수로 이루어진 벡터  $\mathbf{x}$ 와  $c$ 개의 클래스 집합  $\{w_1, w_2, \dots, w_c\}$ 에 대하여  $p(\mathbf{x}|w_j)$ 를  $w_j$ 에 대한  $\mathbf{x}$ 의 클래스-조건부 확률(Class-conditional probability)이라 하고  $P(w_j)$ 를  $w_j$ 의 사전확률(A prior probability)이라 하면 베이저안 정리(Bayesian theorem)에 의해 사후확률(A posterior provability)  $P(w_j|\mathbf{x})$ 를 식 5.2와 같이 구할 수 있다(Duda et al., 2001).

$$P(w_j|\mathbf{x}) = \frac{p(\mathbf{x}|w_j)P(w_j)}{\sum_{k=1}^c p(\mathbf{x}|w_k)P(w_k)} \quad (5.2)$$

사후확률  $P(w_j|\mathbf{x})$ 를 구하는 것은 랜덤변수  $\mathbf{x}$ 에 대하여 임의의 수치로 관측되었을 때 이들이 클래스  $w_j$ 에 얼마나 포함될 지를 확률적으로 예측하는 것과 같

다. 따라서 랜덤변수  $\mathbf{x}$  에 대한  $c$ 개의 클래스  $w_j$ 의 모든 사후확률을  $P(w_j|\mathbf{x})$ 을 이용하여 결정오류를 최소화할 수 있도록 베이저안 결정규칙(Bayesian decision rule)을 적용하면 최적의 결정경계(optimal decision boundary)를 통해 각 클래스를 분류할 수 있다. 이러한 분류방법을 최소오류율 분류(minimum-error-rate classification)라 하며 결정규칙은 다음과 같다(Duda et al., 2001).

$$\text{Decide } w_i \text{ if } P(w_i|\mathbf{x}) > P(w_j|\mathbf{x}) \quad \text{for all } j \neq i$$

이와 같은 결정규칙을 통해 분류문제를 풀기 위해서는 클래스-조건부 확률  $P(w_j|\mathbf{x})$ 와 사전확률  $P(w_j)$ 을 알고 있어야 한다. 만약 클래스-조건부 확률밀도 함수를 가우시안 분포(Gaussian distribution)를 가진다고 가정하면 최우추정법(Maximum likelihood estimation)에 의해 주어진  $n$ 개의 관측데이터로부터 식 5.3~5.4를 적용하면 가우시안 확률밀도함수의 주요 인자인 평균 및 공분산  $\mu, \Sigma$ 의 최적 추정치를 얻을 수 있다(Melsa and Cohn, 1979).

$$\hat{\mu} = \frac{1}{n} \sum_{k=1}^n \mathbf{x}_k \tag{5.3}$$

$$\hat{\Sigma} = \frac{1}{n} \sum_{k=1}^n (\mathbf{x}_k - \hat{\mu})(\mathbf{x}_k - \hat{\mu})^T \tag{5.4}$$

본 연구에서는 요인공간에서 각 클래스의 범위를 효과적으로 지정하기 위해 앞서 기술한 베이저안 결정규칙에 기반한 분류기법을 이용하였다. 이에 앞서 요인분석을 통해 추출된 환경요인의 요인점수를 이용하여 군집분석을 수행하였다. 군집분석은 SPSS를 통해 k-평균 군집분석(k-mean clustering)방법을 이용하였고(Bishop, 2006) 각 샘플 지역을 2개의 군집으로 나누었다. 각각의 군집을 상대적 청정지역과 상대적 오염지역으로 명명하고 이를  $w_1$ 과  $w_2$ 로 표기하였다. 다음으로 각각의 군집  $w_1$ 과  $w_2$ 에 대한 요인점수의 분포를 가우시안 확률밀도 함수로 근사하였다. 이때 가우시안 확률밀도함수의 파라미터인 평균 및 분산은 최우추정법을 통해 추정하였다.

근사된 두 개의 가우시안 확률밀도함수는 요인을 랜덤변수  $f$ 라 할 때 상대적 청정지역에서의 요인의 조건부 확률분포에 대하여  $p(f|w_1)$ 으로 상대적 오염지

역에서의 요인의 조건부 확률분포에 대하여  $p(f|w_2)$ 으로 설명될 수 있다. 여기에  $w_1$  및  $w_2$ 에 대한 사선험률  $P(w_1)$ 과  $P(w_2)$ 는 전체 샘플 수  $N$ 과  $w_1$  및  $w_2$ 의 샘플 수  $n_{w_1}$ 과  $n_{w_2}$ 을 이용하여 각각  $n_{w_1}/N$ 과  $n_{w_2}/N$ 의 관계로부터 추정할 수 있다. 이러한 과정을 통해 베이저안 결정규칙을 적용하여 분류기를 구성하는데 필요한 사후확률  $P(w_1|f)$  및  $P(w_2|f)$ 을 모두 구할 수 있다. 결정오류를 최소화하기 위한 최소오류율 분류기법으로부터 요인 추 상의 한 지점이 바로 상대적 청정지역  $w_1$ 과 상대적 오염지역  $w_2$ 을 구분 짓는 최적의 결정경계가 되며 이 결과는 미생물데이터의 출현패턴에 대한 클래스-조건부 확률밀도함수를 구하는 과정에 직접적으로 이용되었다.

#### 5.2.4. 확률모델을 이용한 미생물 출현패턴 분석

5.2.3절에서는 공통요인의 요인점수가 관측된 지역별로 상대적인 오염정도를 나타내는 것으로 가정하여 요인점수의 분포특성을 가우시안 확률밀도함수로 근사하였고 상대적 청정지역과 상대적 오염지역을 구분할 수 있는 결정경계를 얻었다. 본 논문에서는 이러한 결정경계를 통해 분류된 각 샘플의 클러스터 정보를 기반으로 각각의 미생물에 대한 클래스-조건부 확률모델로 근사하여 오염지역에 대한 미생물의 출현패턴을 분석하였다.

분석의 선행 작업으로 요인-미생물 간의 상관성 분석결과 미생물 데이터의 수치가 이들의 관계를 설명할 수 있는 중요한 의미를 가지고 있지 않음을 확인하였다. 이에 보다 용이한 분석을 위해 미생물 데이터의 값은 1과 0의 두 값을 가지도록 변환하였다. 따라서  $k$ 번째 미생물을  $M_k$ 라 할 때,  $M_k = 1$ 이면, 샘플을 채취한 당시 미생물  $M_k$ 가 발견되었음을 의미하며 이때의 확률을  $P(M_k = 1)$ 로 표현하였다. 또한  $M_k = 0$ 이면, 미생물  $M_k$ 가 발견되지 않았음을 의미하며 이때의 확률을  $P(M_k = 0)$ 라 표현하였다. 이로부터 상대적 청정지역  $w_1$ 에서 미생물이 출현할 확률은  $P(M_k = 1, w_1)$ 으로, 상대적 오염지역  $w_2$ 에서 미생물이 출현할 조건부 확률은  $P(M_k = 1, w_2)$ 라 표현할 수 있다. 여기서 오염정도에 따라 출현하는 임의의 미생물  $M_k$ 에 대하여 요인점수의 분포를 알기위해 랜덤변수  $f$ 를 고려하면 이의 조건부 확률밀도함수는 각각  $p(f|M_k = 1, w_1)$ 과  $p(f|M_k = 1, w_2)$ 이 된다. 즉, 두 확률밀도함수를 구하는 것은 임의의 미생물  $M_k$ 에 대해 출현정보와 클래스 정보를 아는 경우 미생물  $M_k$ 는 요인 축에서 어떠한 분포를 보이는가를 샘플로부터 추정한다는 의미로 볼 수 있다. 한편, 대부분의 미생물의 경우 출현확률  $P(M_k = 1)$ 가 출현하지 않을 확률  $P(M_k = 0)$ 에 비해 현저히 낮기 때문에 미생물이 출현하지 않는 경우에 대한 조건부 확률밀도함수  $p(f|M_k = 0, w_1)$ 과  $p(f|M_k = 0, w_2)$ 은 고려하지 않기로 한다.

조건부 확률 및 사전확률은 주어진 데이터로부터 최우추정법에 의해 추정될 수 있으므로 베이시안 정리로부터 요인점수가 주어진 경우  $w_i$ 지역에서  $M_k$ 가 출현할 사후확률  $P(M_k = 1, w_i|f)$ 는 식 5.5와 같이 계산할 수 있다.



$$P(M_k = 1, w_i|f) = \frac{p(f|M_k = 1, w_i)P(M_k = 1, w_i)}{\sum_{j=1}^2 p(f|M_k = 1, w_j)P(M_k = 1, w_j)} \quad (5.5)$$

각각의 사후확률을 이용하여 미생물의 출현패턴을 분류하기 위해 다음과 같은 기각조건을 두었다. 이는 두 사후확률 모두 임계값  $\theta$ 보다 작을 경우에 해당되며 상대적 청정지역과 상대적 오염지역 모두 미생물  $M_k$ 가 고르게 분포하는 경우를 의미한다.

$$\text{reject if } P(M_k = 1, w_1|f) < \theta \text{ and } P(M_k = 1, w_2|f) < \theta$$

채택된 사후확률에 대해서는 다음의 베이저안 결정규칙을 적용하였다.

$$\text{decide } w_1 \text{ if } P(M_k = 1, w_1|f) > P(M_k = 1, w_2|f)$$

$$\text{decide } w_2 \text{ if } P(M_k = 1, w_1|f) < P(M_k = 1, w_2|f)$$

위의 결정규칙으로부터 미생물의 패턴은 3가지로 분류가 가능하나 기각된 미생물에 대해 두 지역에 고루 분포하며 상관성이 있는 경우 및 무상관인 경우에 해당하는 2가지로 다시 분류하였다. 이때에는 미생물의 수치를 로그변환한 후 요인과의 상관성을 분석하고 각각의 상관계수  $r$ 에 대해  $|r| > 0.4$ ,  $p < 0.001$ 의 경계조건을 이용하여 출현패턴을 분류하였다. 따라서 이러한 과정을 통해 미생물의 출현패턴을 분류하면 다음의 4가지로 정리될 수 있었다. 본 연구에서는 미생물의 출현패턴의 분류결과를 미생물  $M_k$ 를  $P_i$ 로 특성화하였다고 표현하였다.

- (1) 상대적 청정지역에서 주로 출현하는 미생물 : P1
- (2) 상대적 오염지역에서 주로 출현하는 미생물 : P2
- (3) 오염정도와 상관없이 출현하는 미생물 : P3
- (4) 오염정도와 상관하여 출현하는 미생물 : P4

결과적으로 (1)~(3)의 항목에 해당되는 미생물  $M_k$ 는 오염정도와 상관성이 없으나 ( $|r| < 0.4$ ) 청정 지역에서 주로 출현하는 P1, 오염지역에서 주로 출현하는

P2, 청정지역과 오염지역에서 모두 출현하는 P3로 특성화하였다. 또한 청정지역과 오염지역에서 모두 출현하면서 오염정도에 따라 출현밀도가 비례 또는 반비례하는 상관성을 보이는 경우( $|r| > 0.4$ ,  $p < 0.001$ )에는 P4로 특성화하였다.

### 5.2.5. 상대적 중요도 평가를 통한 주요 미생물 추출

미생물을 특성화함으로써 얻을 수 있는 이점은 관측지역의 오염상태를 직관적으로 알 수 있다는 것이다. 그러나 미생물의 종류가 매우 다양하며 특성화에 이용된 미생물의 샘플수가 충분하지 못하다면, 단순히 특성화된 미생물만으로는 정확한 환경평가가 이루어질 수 없다. 따라서 특성화된 미생물 중에서 환경평가에 상대적으로 많은 영향을 미치는 미생물을 선택하는 것은 의미가 있는 일이다. 이를 위한 기본적인 접근방법은 특성화된 미생물의 출현빈도가 높은 미생물 순으로 순위를 매기고 상위그룹에 있는 미생물을 추출하는 것이다. 그러나 본 연구에서 다루는 미생물 데이터는 관측된 샘플의 개수가 매우 적기 때문에 단순히 빈도수만으로 중요 미생물을 선별하는 것은 부적합하다. 또한 미생물의 종류가 매우 다양하기 때문에 동일한 환경조건에서 각기 상반된 특성을 보이는 미생물들이 존재 가능하며 이러한 특성을 가지는 미생물을 어떻게 처리할 것인지 등에 대한 효과적인 방법론이 필요하다. 즉 미생물 개개의 특성과 같이 상대적 특성 등을 종합적으로 고려해야만 올바른 환경평가가 가능함으로 다음의 사항을 특히 고려하여 미생물을 선별하고자 하였다. 첫째, 선별된 미생물을 이용하여 오염정도를 판정하기 위한 최소한의 분류기를 구성할 수 있어야 한다. 둘째, 주요 미생물은 각 미생물간의 상대평가를 통해 선별되어야 하며 이들은 상대적으로 우수한 예측성능을 가져야 한다.

이에 본 연구에서는 주요 미생물들을 선정하기 위해 미생물의 상대적인 중요도를 측정할 수 있는 평가시스템을 구성하였다. 이 평가시스템에는 오염도 평가를 위해 샘플 데이터로부터 학습이 가능한 기로서의 분류기가 포함되도록 구성하였다. 분류기의 학습이 완료되어 미생물의 출현 밀도에 대한 입력이 주어지면 분류기는  $w_1$  및  $w_2$ 의 두 상태에 대응되는 값을 출력한다. 각 변수의 상대적 중요도의 측정은 각각의 미생물 변수에 대한 특성들이 분류기를 통한 예측결과에 얼마나 잘 반영이 되는지를 측정하는 것으로 이를 반복적인 모의실험(simulation)으로부터 계산될 수 있도록 시스템을 구성하였다. 이러한 시스템의 특성상 평가시스템에 포함된 분류기는 빠른 학습이 가능하며 분류문제에 적합한 확률신경망(PNN: Probabilistic Neural Network)을 이용하였다.

확률신경회로망은 RBF(radial basis function)과 베이지안 분류기가 신경망의 형태로 구성되어 있으며 학습의 과정이 최소제곱 알고리즘을 통해 가장 가깝게

인접하고 있는 데이터들을 같은 클래스의 패턴으로 분류하도록 하는 가중치들이 조정됨으로써 학습과정이 일반적인 MLP에 비해 빠른 장점이 있다(Spekt, 1990). 확률신경망은 Fig. 5.2와 같이 RB층(radial basis layer)과 경쟁층(competitive layer)로 이루어져 있다. RB층의 활성화 함수는 가우시안 함수를 통해 입력벡터의 일정 영역 내에 존재하는 입력 값들만을 채택하도록 하는 구조를 가진다. Fig. 5.3은 주어진 입력 데이터에 대한 채택 및 기각에 대한 개념을 나타낸 것으로서 너비 파라미터  $\sigma$ 의 조정을 통해 기각영역을 결정한다.

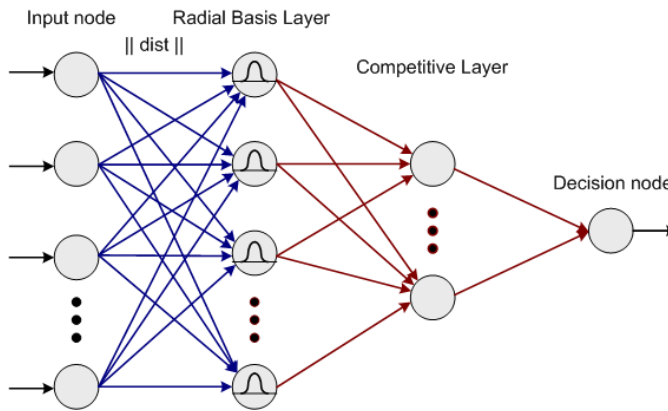


Fig. 5.2. Schematic diagram probabilistic neural networks.

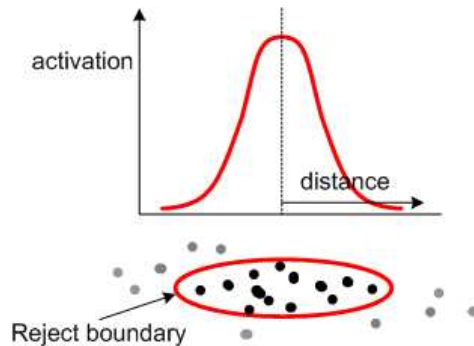


Fig. 5.3. Concept of reject boundary in input space using a radial basis function with gaussian kernel.

이를 통해 임의의 입력벡터가 채택영역의 중심근방에 위치하면 이와 관련된 은닉뉴런은 활성화되며 두 개의 채택영역 중심의 사이에 위치하고 채택 폭  $\sigma$  내에 존재한다면 이때의 뉴런들은 부분적으로 활성화된다. 경쟁 층에서는 RB 층에서 얻어진 각 입력데이터들의 클래스 소속에 대한 확률밀도함수를 결합한 후 최대의 확률을 가지는 뉴런을 선택하는 역할을 수행하며 식 5.6과 같은 가우스 가중함수를 통해 확률밀도함수를 추정하게 된다(이현엽, 문경일 1999).

$$g(x) = \frac{1}{n} \sum_{i=1}^n \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (5.6)$$

여기서  $n$ 은 사례의 개수이고  $x_i$ 는 클래스 내의 특정 사례를 나타내며  $x$ 는 입력,  $\sigma$ 는 너비파라미터를 나타낸다. 최종적으로 결정노드는 이 확률이 최대가 되는 노드 혹은 클래스를 선택하여 출력한다.

본 연구에서는 확률신경망을 44개의 미생물 변수에 대한 입력벡터와 2개의 클래스 정보를 출력으로 하도록 구성하였다. 또한 상대적 중요도 평가가 반복적인 모의실험을 통해 이루어지기 때문에 각 단계마다 확률신경망의 학습 및 평가를 위해 미생물 데이터를 무작위로 추출하여 80:20의 비율을 가지도록 배분하였다. 이는 주어진 학습데이터에 최적화된 분류기나 상대적 중요도의 결과를 얻는 것 보다는 확률적으로 신뢰되는 결과를 얻기 위한 평가 전략이라 할 수 있다. 따라서 상대적 중요도는 학습과 평가를 반복됨에 따라 확률적으로 최적의 수치 값으로 수렴하며 이러한 결과를 통해 우수한 성능의 오염예측시스템을 구성하는 데 필요한 예측능력이 높은 중요한 변수의 선별이 가능하다. Fig. 5.4는 이러한 사항을 고려하여 설계된 상대적 중요도 평가시스템의 구조를 나타내었다.

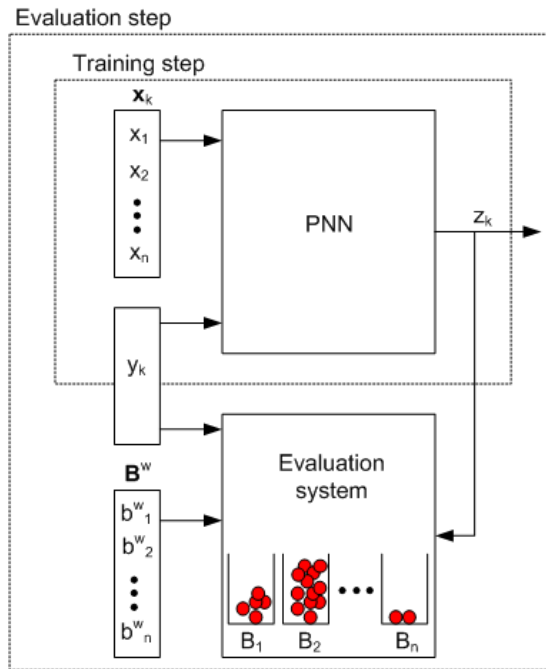


Fig. 5.4. Structure of the proposed system for measuring the relative importance among microorganisms.

상대적 중요도 평가 알고리즘은 다음과 같은 단계를 반복함으로써 각 미생물의 상대적 중요도를 측정된 결과를 확률 값으로 반환한다. 우선 분류기의 학습을 수행하기 위해 전체의 80%에 해당하는 데이터를 무작위로 추출하여 학습데이터로 사용하며 나머지 20%를 평가를 위한 데이터로 분류한다. 확률신경망 분류기는 입력 값과 목표 값을 가지는 데이터 집합으로부터 학습하는 지도학습(Supervised learning)방법에 의해 학습되는데 본 연구에서는 입력벡터를 각 미생물 데이터의 로그변환 수치로 두었고 목표 값은 5.2.3절에서 기술한 방법으로부터 얻은 클러스터 정보를 가지도록 하였다. 확률신경망의 연결강도에 대한 학습이 완료되면 PNN 분류기는 미생물을 이용하여 오염여부를 판정하는 예측시스템으로서 작동 하게 된다. 이때 5.2.4절로부터 얻어진 각 미생물의 특성치를 이용하여 상대적인 평가를 수행한다. 테스트 데이터로부터 분류기가 올바른 예측결과를 출력할 경우에 한하여, 이때 이용된 미생물의 특성치가 예측결과와 같은 값을 가지는 경우 그 변수에는 점수가 부여된다. 이러한 방법으로 주어진

테스트 데이터를 모두 이용하여 평가를 수행하면 각각의 변수들에 대한 점수가 할당되는데, 이때 높은 점수를 가진 변수는 미생물의 특성이 잘 반영된 예측 성능이 좋은 변수임을 나타낸다. 학습 및 평가의 한 과정이 끝나면 각각의 변수들에 대한 예측성능은 식 5.7과 같이 계산된다.

$$R_i = \sum_k h_k^2 E(b^{w_k} = w_i) \quad (5.7)$$

여기서  $h_k$ 는 k미생물에 대한 예측확률을 나타내며,  $E(b^{w_k} = w_i)$ 는 클래스 소속 함수로  $w_k$ 로 특성화된 k 미생물 변수들이 목표상태 및 출력상태인  $w_i$ 와 같을 경우 1을, 아닐 경우에는 0을 반환한다. 알고리즘은 이러한 학습 및 평가를 일정횟수만큼 반복하여 수행하게 되는데 반복평가가 모두 종료되면 상대적 중요도는 식 5.8과 같이 계산되며 이는 각 평가단계에서 구한 예측성능의 평균을 나타낸다.

$$RI = \frac{1}{N} \sum_{i=1}^N R_i \quad (5.8)$$

## 5.2.6. 미생물 지수화

상대적 중요도가 높은 순으로 선별된 미생물의 정보는 오염정도를 평가하는데 중요한 정보를 제공하며 일부 미생물은 출현 그 자체만으로도 수질의 오염 여부를 판정할 수도 있다. 이를 위해서는 오랜 시간의 관측 데이터에 대한 통계적 결과가 요구되지만 본 연구에서 제시한 접근방법은 주요한 미생물을 선정하는 방법론적 측면에서 의미가 있다. 일단 미생물이 선정이 된 경우 이들을 이용한 다양한 형태의 예측시스템을 구축할 수 있겠으나 본 연구에서는 오염정도를 대표하는 지표로서의 미생물들을 열거하고 이들이 지닌 고유의 영향력을 지수로 부여하여 최종적으로는 보다 직관적인 환경평가를 보다 쉬운 방법으로 수행하고자 하는 것을 목적으로 한다. 우선 미생물 지수를 이용하여 환경평가를 수행할 때 평가의 정도나 계산상의 편의성 등을 고려하여 10개의 등급으로 나눈다. 이렇게 할당된 미생물 지수는 BMWP 등의 다른 환경평가를 지수와 유사한 의미로서 해석될 수 있는데 10등급은 상대적으로 청정지역을, 1등급은 상대적으로 오염지역을 나타낸다. 선정된 미생물은 상대적 중요도에 따라 간단한 스케일 함수에 의해 등급별로 분류되며 각각 지수점수를 1에서 10점사이의 정수 값을 가지도록 하였다. 이렇게 선정된 미생물 지수를 환경평가에 활용하는 방법은 다음과 같다. 임의의 지역에서 PCR-DGGE분석을 통해 관측된 미생물 중 지수가 부여된 미생물만을 추출한다. 지수가 부여된 미생물은 각각 출현 밀도에 관련된 수치 값이 기록되어 있으며 이 값은 각각의 미생물에 대한 가중치로서 이용된다. 최종 평가를 위해 각 미생물의 지수점수와 가중치를 이용한 가중 합을 계산하는데 이는 1에서 10 사이의 범위를 가지는 실수가 되며 이 수치는 지수점수의 의미와 동일한 의미를 가진다.

미생물에 오염도에 대한 지수로서 미생물을 나타내기 위해 앞 절에서 구한 상대적 중요도를 이용하여 청정지역 및 오염지역에 대한 미생물들의 상대적 중요도를 반대의 부호를 지니도록 한 후 요인 축 상에서 중심점으로부터 최대  $\alpha$ 의 거리를 두도록 mapping하였다.  $\alpha$ 는 요인의 표준편차의 2배인(미생물 데이터가 약 98%가 밀집하는 경계) 2를 부여하였는데 이는 본 연구에서 개발하고자 하는 미생물 지수의 최대값 및 최소값이 요인 축 상에서의 위치가 얼마나 되는지를 쉽게 추측할 수 있다는 이점이 있다. 여기서  $b^w(l)$ 은 미생물의 특성을 구분하는 부호결정을 위한 함수로 청정지역인  $w_1$ 에 대해서는 -1을 오염지역인  $w_2$



에 대해서는 1을 반환한다.  $RI_l$  는 상대적 중요도를 나타내며 최종적으로  $l$ 번째 미생물의 지수는 식은 식 5.9와 같다.

$$I_l = \alpha b^w(l) RI_l \quad (5.9)$$

선택된 미생물들에 한해 지수가 부여되면 식 5.10에 의해 주어진 미생물데이터와 미생물 지수를 이용하여 환경오염정도를 평가할 수 있다.

$$J = \sum_{i=1}^{10} I_i S_i \quad (5.10)$$

여기서  $S_i$ 는  $i$ 그룹에 속하는 유효한  $k$ 개의 미생물에 대한 밀도의 총합을 전체 밀도의 총합으로 나눠준 것으로 결국  $S_i$  는 10개의 그룹에 대한  $i$ 그룹의 미생물에 대한 확률값으로  $i$ 그룹에 대한 가중치를 나타내며 식 5.11과 같이 계산된다.

$$S_i = \frac{\sum_{j=1}^k B_j^i}{\sum_{i=1}^{10} \sum_{j=1}^k B_j^i} \quad (5.11)$$

### 5.3. 결과 및 고찰

#### 5.3.1. 요인분석에 의한 환경요인의 추출 및 환경변수와의 관계

수집된 13개의 환경변수에 대한 환경데이터를 이용하여 환경평가에 활용될 수 있는 주요 정보를 추출함에 있어 보다 신뢰되는 결과를 얻기 위해 요인분석 기법을 이용하여 분석을 실시하였다. 요인분석은 다수의 변수들 간의 상관관계를 분석하여 변수들의 바탕을 이루는 공통차원들로서 이 변수들을 설명하는 통계기법이다. 따라서 본 연구에서 분석하고자 하는 환경데이터와 같이 시료수에 비해 고차원의 변수로 구성된 데이터에서 저차원의 주요 요인을 추출하여 이를 이용하여 원래의 데이터 패턴을 나타낼 수 있기 때문에 보다 효과적인 분석이 가능하다. 요인분석에 앞서 환경데이터를 로그공간으로 변환한 후 정규화를 수행하였다. 이는 관측 값들의 지수적 분포로 인해 편향된 분석결과를 지양하기 위함이며 요인분석을 위한 데이터는 평균이 0이고 분산이 1인 정규분포를 가져야 하는 요인분석의 기본가정 때문이다. 전처리 과정을 수행한 후 데이터에 포함된 주요 요인의 속성을 알아보기 위해 13개의 환경변수 전체를 대상으로 요인분석을 수행하였다. 요인의 추출방법은 주성분분석을 이용하였고 추출된 요인을 베리맥스(Varimax) 방법으로 회전시켜 각 요인에 대해 변수의 영향력을 쉽게 파악할 수 있도록 하였다. 13개의 환경변수를 이용하여 요인분석을 수행하게 되면 최대 13개의 요인을 얻을 수 있는데 이때 고유값이 1이상인 요인만을 추출한 결과 4개의 주요 요인을 얻을 수 있었다. 또한 각 요인별로 회전된 요인 값의 절대치가 0.5 이상인 경우에 해당되는 변수를 그 요인에 의해 설명가능하다고 간주하였다. 이러한 조건으로부터 제 1요인에는 velocity, turbidity, BOD, T-P, T-N 및  $\text{NO}_3^-$ -N가 다른 환경변수에 비해 두드러진 결과를 보였다. 제 2요인에는 Depth 와 TOC, 제 3요인에는 Conductivity, 제 4요인에는 TKN이 두드러진 수치를 보였다. Table 5.1은 환경변수 13개를 대상으로 요인분석 결과를 나타내었다. 여기서 추출된 4개의 주요요인에 대해서 중요한 변수로 간주되는 항목은 굵은 글씨로 구별되어 있다. 보다 가시적인 분석을 위해 3개의 주요요인을 축으로 하여 각 변수들의 공간적인 위치를 표시하면 Fig. 5.5와 같다. 그러나 이들의 위치가 다소 산만하게 분포되어 있음을 확인할 수 있었다.

Table 5.1. Factor loadings of 4 principal factors extracted by factor analysis for all of the environment variables tested.

| Environmental parameter         | Factor       |        |        |        |
|---------------------------------|--------------|--------|--------|--------|
|                                 | 1            | 2      | 3      | 4      |
| pH                              | -0.134       | 0.254  | -0.140 | 0.190  |
| Temperature                     | 0.007        | -0.009 | 0.011  | 0.390  |
| Conductivity                    | 0.063        | -0.020 | 0.995  | 0.062  |
| Velocity                        | <b>0.572</b> | 0.416  | -0.283 | 0.049  |
| Turbidity                       | <b>0.649</b> | -0.211 | 0.071  | 0.063  |
| Depth                           | 0.008        | 0.665  | 0.163  | 0.048  |
| BOD                             | <b>0.540</b> | -0.238 | 0.221  | 0.150  |
| COD                             | 0.043        | 0.007  | -0.029 | 0.550  |
| TOC                             | 0.159        | -0.741 | 0.218  | 0.269  |
| T-P                             | <b>0.746</b> | -0.209 | 0.188  | -0.015 |
| TKN                             | 0.081        | -0.053 | 0.090  | 0.604  |
| T-N                             | <b>0.912</b> | 0.083  | -0.056 | 0.131  |
| NO <sub>3</sub> <sup>-</sup> -N | <b>0.770</b> | 0.490  | -0.207 | -0.217 |

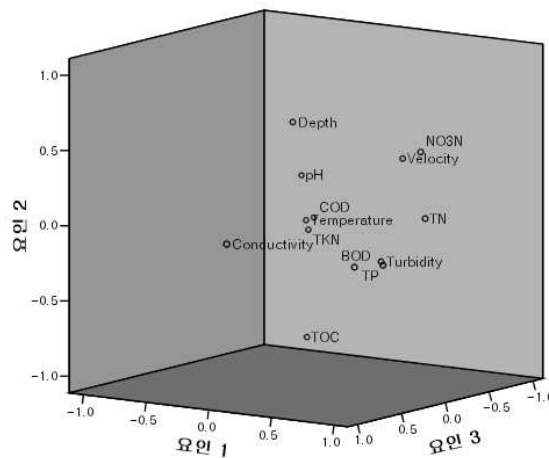


Fig. 5.5. A result of mapping the environment variables into principal factor space.

이는 하나의 변수가 하나의 요인으로 모두 설명되는 것이 아니기 때문에 설명 제 1요인에 높은 요인 값을 가졌다 하더라도 velocity나 NO<sub>3</sub><sup>-</sup>-N 등과 같이 다른 요인에도 연관되어 있기 때문이다. 따라서 공통성 또는 상관성이 작거나 불필요한 변수들을 제거하고 요인분석을 수행하는 것을 반복함으로써 분석하고자 하는 데이터에서 주요요인을 효과적으로 추출할 수 있다(이, 2001). 제 1요인은 많은 변수들의 특성을 가장 잘 반영하는 요인이므로 제 1요인으로 구분되지 않고 상관성이 낮은 변수를 차례로 제거하여 하나의 요인으로 설명 가능한 변수들만 추출하였다. 13개의 항목에 대해 인자분석을 실시하여 상관관계를 이용해 여분의 항목을 제외시켜 T-P, T-N, NO<sub>3</sub><sup>-</sup>-N, BOD, Turbidity가 하나의 요인으로 추출되었고 Table 5.2는 추출된 변수들을 이용하여 요인분석 수행 결과로부터 산출된 요인적재량(factor loading)을 나타낸다. 요인 적재량은 추출된 요인과 식 5.12의 선형결합을 통해 특정 변수들의 정규화 값으로 변환가능하다.

$$z_i = l_i F + \epsilon_i \quad (i = 1, \dots, 5) \tag{5.12}$$

요인분석을 통해 추출된 변수들은 각각 수질오염을 평가하는데 중요한 인자로 판명되어 있다. 따라서 이들 변수를 대표하는 공통요인 역시 수질오염 정도를 평가하는데 활용이 가능하다. 따라서 미생물과 각 환경변수들의 상관성을 분석하기 위해 5 가지의 환경변수 모두를 이용하는 것이 아니라 요인만을 이용한다고 하더라도 신뢰되는 결과를 얻을 수 있다.

Table 5.2. Factor loadings for the 5 environment variables selected by factor analysis.

| Environmental parameter         | Factor |
|---------------------------------|--------|
|                                 | 1      |
| T-P                             | 0.694  |
| T-N                             | 0.957  |
| NO <sub>3</sub> <sup>-</sup> -N | 0.740  |
| BOD                             | 0.531  |
| Turbidity                       | 0.608  |

식 5.12는 요인분석을 통해 얻은 요인과 환경변수와의 선형관계식을 나타내지만 정규화된 환경변수  $z_i$ 가 가지는 고유의 변수  $\epsilon_i$ 의 차이만큼 오차가 존재한다. 따라서  $\epsilon_i$ 까지 모두 고려하여 회귀분석을 수행하면 요인과 환경변수와의 보다 정확한 상관관계를 얻을 수 있다. 따라서 요인과 미생물의 관계가 일단 규명되면 요인과 각 환경변수들의 관계식을 회귀분석을 통해 간단히 얻을 수 있으므로 최종적으로 환경변수와 미생물의 관계도 알 수 있다. Fig 5.6은 요인분석을 통해 얻은 요인점수와 각 환경변수의 정규화 값과의 관계를 도표로 나타내었다. 도표를 통해 알 수 있듯이 요인과 환경변수의 상관관계는 높은 양의 관계를 가지고 있음을 확인하였다. 특히 요인분석결과의 요인적재량을 통해 확인할 수 있듯이 요인은 T-N과 매우 흡사한 분포패턴을 가지고 있음을 알 수 있었다. 반면 T-P 및 BOD와 요인의 관계는 다소 산만한 데이터 분포를 가지고 있으나 환경평가를 위한 주요 변수로서 이용될 가치는 충분히 있다고 사료된다. 요인과 환경변수들 간의 회귀분석을 수행한 결과로 이들의 구체적인 관계를 나타내는 각 계수들은 Table 5.3과 같이 계산되었다. 회귀분석은 요인을 독립변수로 두고 환경변수의 정규화 변수를 각각의 종속변수로 간주하여 시행하였다.  $i$ 번째 환경변수에 대하여  $\beta_i$  와  $\alpha_i$  는 각각 회귀분석을 통해 분석된 요인  $F$ 과 정규화변수  $z_i$ 를 관계시키는 기울기 및 절편을 나타내며  $\mu_i$ 와  $\sigma_i$ 는 환경변수의 정규화과정에서 얻어진 평균 및 표준편차를 나타낸다.

Table 5.3. Coefficients for mapping factor-space into variable-space for the selected environment variables.

| Environmental variables         | $\beta$  | $\alpha$  | $\mu$    | $\sigma$ |
|---------------------------------|----------|-----------|----------|----------|
| Turbidity                       | 0.796081 | 0.017478  | 2.398268 | 1.305596 |
| BOD                             | 0.662111 | -0.023966 | 2.280118 | 1.179129 |
| T-P                             | 0.570513 | -0.107947 | 0.757462 | 0.724734 |
| T-N                             | 0.963254 | 0.001772  | 2.924338 | 1.643697 |
| NO <sub>3</sub> <sup>-</sup> -N | 0.834246 | 0.123240  | 3.023255 | 1.326330 |

이러한 계수들을 이용하면 임의의 요인점수를 환경변수  $x_i$ 의 값으로 변환시킬 수 있으며 그 관계식은 식 5.13과 같다.

$$\begin{aligned} z_i &= \beta_i F + \alpha_i \\ x_i &= \sigma_i z_i + \mu_i \end{aligned} \tag{5.13}$$

Fig 5.6 ~ Fig 5.10은 5 개의 환경변수들에 대한 회귀분석의 결과를 요인점수에 적용하여 도시한 것이며 요인분석으로부터 얻은 식 5.12의 관계를 비교 검토하기 위해 회귀분석 결과와 함께 도시하였다. 여기서 좌측 상단의 식은 회귀분석의 결과로부터 주어진 식 5.12의 관계식을 나타내며 우측하단의 식은 요인분석의 결과로부터 주어진 식 5.13의 상위 관계식을 나타낸다. 대부분의 경우 두 직선이 다소 차이를 보였으나 T-N의 경우에는 두 직선이 거의 일치하는 것으로 분석되었다. 이는 환경변수 T-N의 고유 분산성분이 나머지 4개의 환경변수들의 공통 분산성분들을 대표할 수 있을 정도로 잘 설명할 수 있기 때문에 요인분석 결과로 추출된 공통요인에 대해서 T-N의 요인적재량이 0.957로 매우 높았고, 이러한 이유로 회귀분석 결과와 요인추출결과로 얻어진 두 직선이 거의 일치하는 것으로 사료된다. 또한 나머지 4개의 환경변수와 요인과의 관계에서 회귀분석과 요인분석의 결과가 차이를 보이는 것은 각 변수에 대한  $\epsilon_i$ 의 고유 분산성분에 의한 영향인 것으로 사료되며, 이는 각 변수에 대한 전체 분산 성분 중 PCA에 의해서 공통 분산성분만이 요인추출에 이용되고 남겨진 고유 분산성분으로 이는 공통요인에 대해서 설명되지 않는 성분으로 사료된다. 그러나 회귀분석은 요인과 각 환경변수간의 분산성분으로 관계되어지는 것이기 때문에 이러한 차이를 보이는 것으로 판단된다.

한편 시료를 채취한 각 지역에서 요인점수와 환경변수들 간의 지역별 환경상태의 경향을 파악하기 위해 각 지역을 요인점수에 따라 두 그룹으로 임의로 나누어 분석하였다. 분석결과 샘플링 지역별로 환경변수 또는 요인의 값에 따라 몇 개의 군집으로 분류가 가능함을 알 수 있었다. 이는 요인분석을 통해 추출된 요인이 수질상태가 다른 지역들의 각 환경변수들의 측정값들을 잘 반영하는 것이며 이러한 관계를 이용하여 특정지역에 대한 환경상태를 요인점수만을 통해서도 평가가 가능하다는 결론을 얻을 수 있었다.

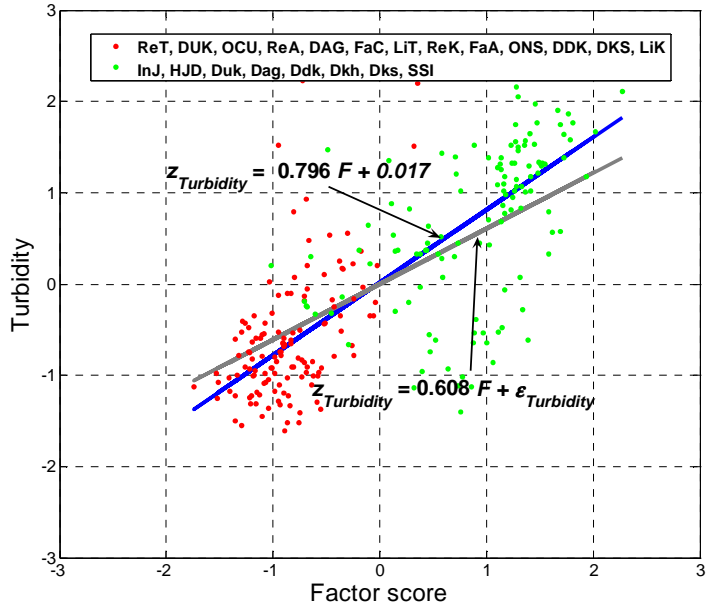


Fig. 5.6 A comparative analysis to reveal the relation between common factor and turbidity variable through regression and factor analysis.

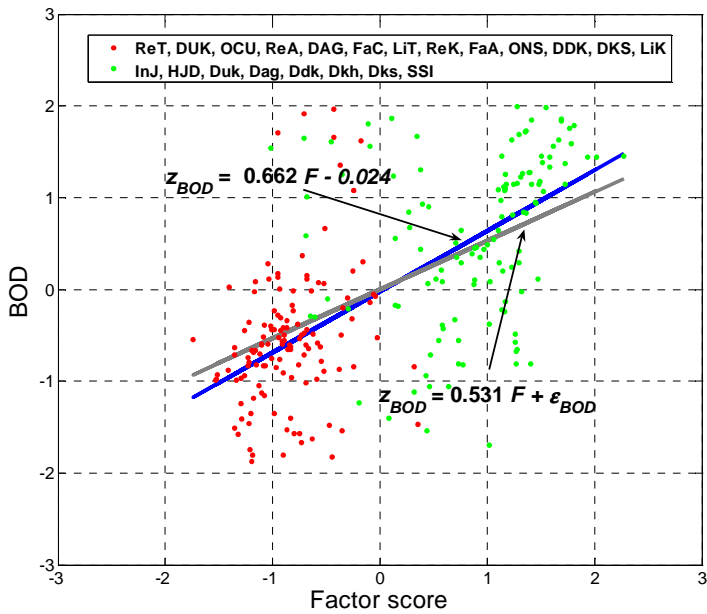


Fig. 5.7 A comparative analysis to reveal the relation between common factor and BOD variable through regression and factor analysis.

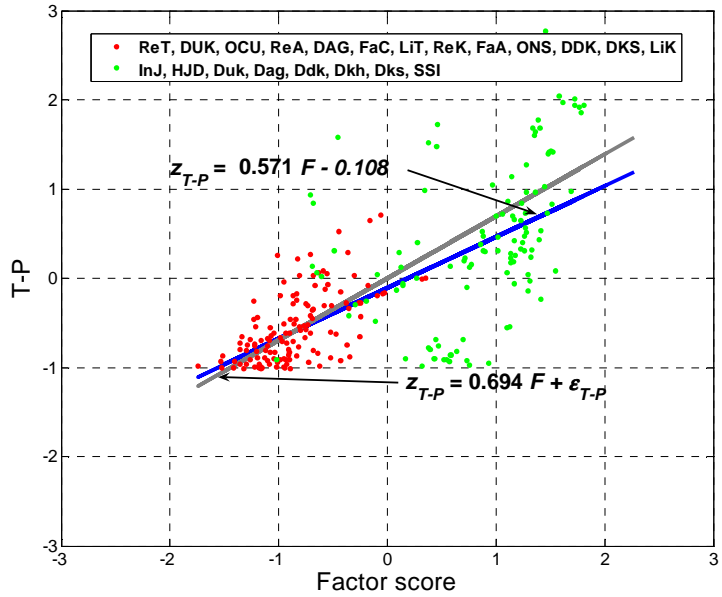


Fig. 5.8 A comparative analysis to reveal the relation between common factor and T-P variable through regression and factor analysis.

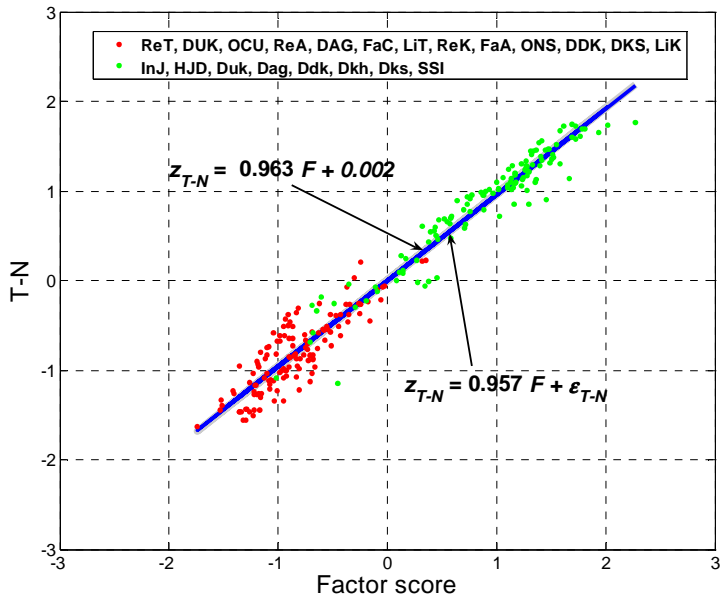


Fig. 5.9 A comparative analysis to reveal the relation between common factor and T-N variable through regression and factor analysis.



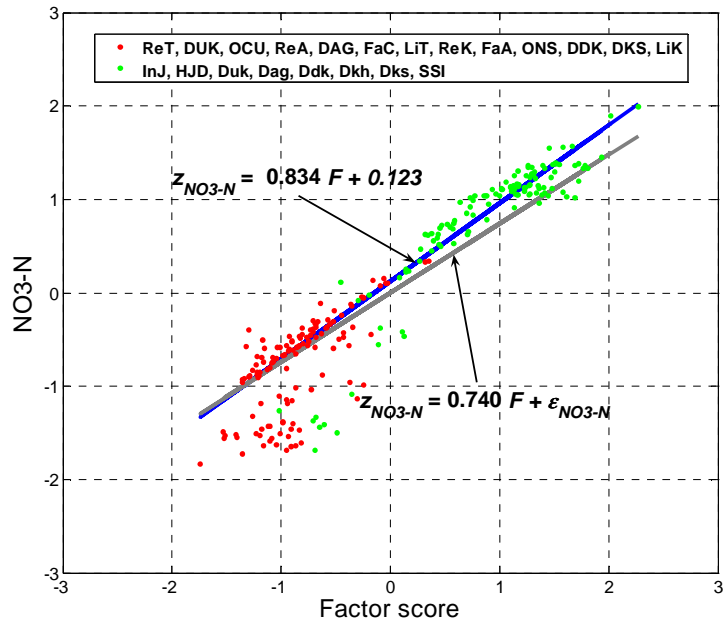


Fig. 5.10 A comparative analysis to reveal the relations between common factor and NO<sub>3</sub><sup>-</sup>-N variable through regression and factor analysis.

### 5.3.2. 요인분석에 의해 추출된 환경요인과 미생물 군집과의 관계

요인분석을 통해 추출된 환경요인은 5개 환경변수를 대표하는 공통요인으로 수질의 상태를 예측하는데 유용한 정보가 될 수 있음을 확인하였다. 이를 통해 추출된 환경요인은 5개의 환경변수를 대신하여 미생물과의 관계를 규명하는데 있어 활용될 있다. 따라서 환경과 미생물 군집간의 관계를 분석하기 위하여 이들 간의 상관성을 우선적으로 분석하였다. 구체적으로 미생물의 값이 기록되지 않은 경우를 모두 제외한 경우에 대해 추출된 요인과 115종의 미생물 자료에 대한 상관관계를 분석하였다. 그러나 상관성 분석결과 요인점수에 대한 상관성이 낮은 경우가 대부분이었고 상관성이 높다하더라도 유의수준이 높아 신뢰할 만한 상관관계를 설명하기 어려웠다. 이는 미생물 변수가 환경데이터와는 달리 상당부분이 관측이 불가하였으며 이로 인해 보다 면밀한 상관여부를 내리는 것이 어렵기 때문인 것으로 사료된다. 115종의 미생물 중 상관성이 높은 순으로 정렬한 후 상위 3개의 미생물변수와 요인점수와의 관계를 Fig. 5.11 ~ Fig. 5.13에 나타내었다. 요인점수와의 상관성이 상대적으로 높은 3개의 미생물은 각각 *Acinetobacter* sp., *Sphingomonas* sp., *Pseudomonas* sp. 로 밝혀졌다. 그러나 상관성이 가장 높은 *Acinetobacter* sp.의 경우 상관계수가 0.43정도이며 나머지 두 종에 대해서는 0.3이하의 상관계수를 나타내었다. 이때 *Acinetobacter* sp.의 관측확률은 39%로 모든 미생물 중에서 가장 많이 관측되어진 것으로 판명되었다. 이러한 결과로부터 변수들 간의 관계를 알기 위해 보편적으로 이용되는 상관성분석이나 회귀분석법은 환경과 미생물의 관계를 분석함에 있어 적합하지 않음을 확인할 수 있었다.

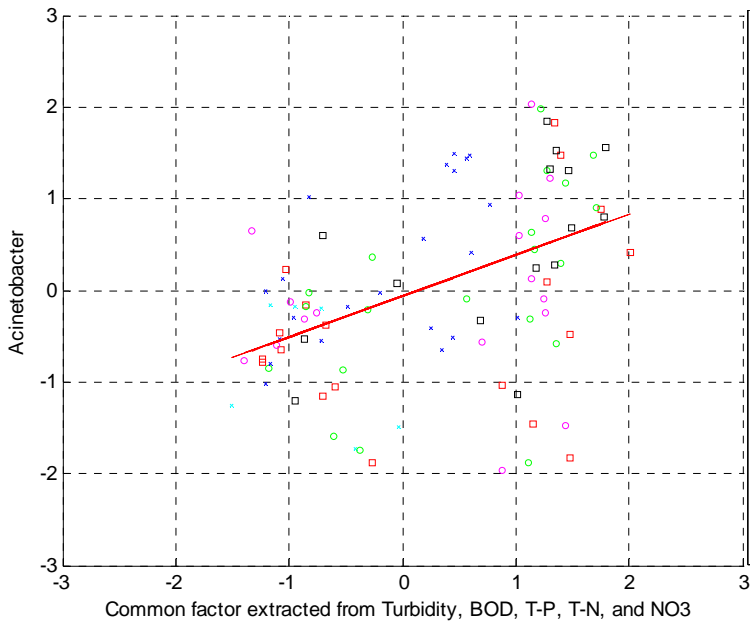


Fig. 5.11. Correlation between the common factor and *Acinetobacter* sp. ( $r = 0.431010$ ,  $p = 0.000008$ ,  $P(n) = 0.391304$ ).

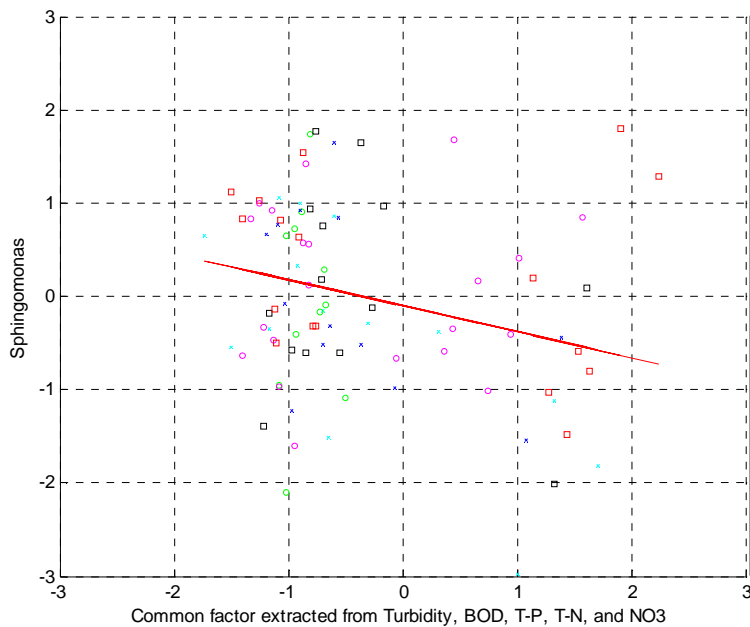


Fig. 5.12. Correlation between the common factor and *Spingomonas* sp. ( $r = -0.234947$ ,  $p = 0.025811$ ,  $P(n) = 0.355731$ ).

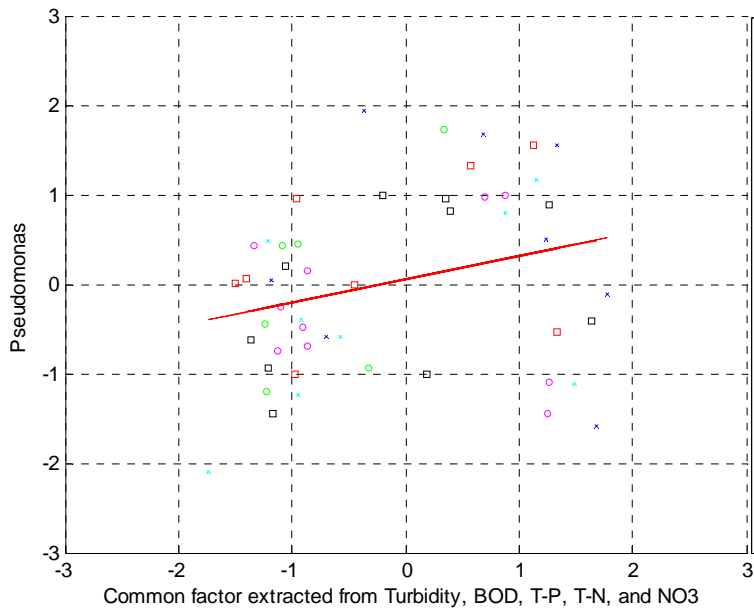


Fig. 5.13. Correlation between the common factor and *Pseudomonas* sp. ( $r = 0.257251$ ,  $p = 0.071307$ ,  $P(n) = 0.197628$ ).

### 5.3.3. 환경요인에 대한 확률모델 및 오염지역 분류

5.3.2절에서 언급한 바와 같이 미생물과 오염 요인간의 관계를 회귀분석을 통해 규명하는 것은 쉽지 않았다. 이것은 관측된 미생물의 종류가 많은 반면 상대적으로 개개의 미생물 데이터는 매우 적었다. 또한 주어진 요인에 대한 데이터의 분포가 매우 산만하고 각 미생물간에 서로 독립적인 분포패턴을 가지기 때문이다. 이에 본 연구에서는 미생물데이터의 수치적 의미보다는 어떠한 조건에서 출현하며 그때의 출현 패턴이 어떠한지를 확률적 분석방법을 통해 규명하고자 하였다. 더불어 이를 이용하여 미생물의 지수를 부여하거나 특정미생물이 출현하는 환경의 조건의 예측이 가능한지를 타진하였다.

분석을 용이하게 하기 위하여 우선 요인분석을 통해 추출된 환경요인의 요인점수를 군집분석을 수행하였다. k-평균 군집분석을 시행한 결과 대부분의 경우 각 샘플링 지역별로 명확한 군집결과를 얻을 수 있었으나 DKS, LiK, InJ, HJD 등의 경계지점에 속하는 지역이나 일부 노이즈가 포함된 지역은 클러스터 정보가 두 지역 모두를 포함하는 경우도 발견되었다. 이러한 경우 각 시료채취 지역에서 보다 많은 클러스터 정보를 가지는 군집으로 분류되도록 하였다. 분류된 두 개의 군집은 각각 상대적 청정지역  $w_1$ 과 상대적 오염지역  $w_2$ 라 명하였다. 이는 요인점수를 기준으로 요인점수가 낮은 경우는 5대 환경변수의 실측값이 상대적으로 낮은 경우로 상대적 청정지역이며, 요인점수가 높은 경우는 환경변수의 실측값이 높은 상대적 오염지역으로 구분하였다. 각 지역별로 군집화를 수행한 결과와 각 지역별 평균 요인점수는 Table 5.4와 같다. DKS, LiK, InJ, HJD의 지역은 요인점수가 결정경계 근처에 분포하였으나 DKS와 LiK은 상대적 청정지역에, InJ와 HJD는 상대적 오염지역에 분류되었음을 알 수 있다. 또한 각 지역별로 요인점수를 비교하면 청정지역에서는 상대적으로 작은 요인점수를, 오염지역은 큰 요인점수를 가지고 있음을 알 수 있다.

Table 5.4. Two-category classification based on probabilistic model and average scores of factors in the sampling sites.

| <i>Group</i> | <i>Site</i> | <i>Factor score</i> |
|--------------|-------------|---------------------|
| $w_1$        | ReT         | -1.157154           |
|              | DUK         | -1.141357           |
|              | OCU         | -1.076308           |
|              | ReA         | -.964738            |
|              | DAG         | -.927401            |
|              | FaC         | -.890950            |
|              | LiT         | -.848284            |
|              | ReK         | -.817104            |
|              | FaA         | -.742444            |
|              | ONS         | -.724743            |
|              | DDK         | -.637922            |
|              | DKS         | -.625321            |
|              | LiK         | -.592938            |
| $w_2$        | InJ         | -.401249            |
|              | HJD         | .105592             |
|              | Duk         | .496651             |
|              | Dag         | 1.040630            |
|              | Ddk         | 1.260975            |
|              | Dkh         | 1.348111            |
|              | Dks         | 1.355037            |
|              | SSI         | 1.417262            |

오염여부에 따라 두 개의 군집으로 분류된 요인점수의 분포를 보다 정확히 분석하기 위해 이들의 확률분포 분석을 시행하였다. 확률분포 분석의 목적은 요인점수가 주어진 경우 두 군집 중 어느 군집에 속할지를 확률값으로 알기 위한 것으로 이는 사후확률  $P(w_i|f)$ 을 구하여 계산할 수 있다. 5.2.3절에서 설명한 바와 같이 사후확률  $P(w_i|f)$ 을 구하기 위해서는 클래스-조건부 확률밀도함수  $p(f|w_i)$ 와 사전확률  $P(w_i)$ 를 추정해야 한다. 이를 위해 최우추정법을 통해 요인점수의 분포를 가우시안 확률밀도함수로 근사하였다.  $W_i$ 를  $w_i$ 에 해당되는 환경요인 데이터의 인덱스로 구성된 집합이라고 하면  $i$ 의 인덱스를 가지는 군집에 대한  $p(f|w_i)$ 은 1차 가우시안 확률밀도함수로서 파라미터인 평균  $\hat{\mu}_i$  및 분산  $\hat{\sigma}_i$ 은 식 5.3~5.4를 적용하여 식 5.14~5.15를 통해 추정되었다.

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{k \in W_i} f_k \quad (5.14)$$

$$\hat{\sigma}_i = \frac{1}{n_i} \sum_{k \in W_i} (f_k - \hat{\mu}_i)^2 \quad (5.15)$$

여기서  $n_i$ 는 집합  $W_i$ 에 포함된 데이터의 총 수를 나타내며  $f_k$ 는  $k$ 번째 데이터의 요인점수를 나타낸다. 사전확률  $P(w_i)$ 는 5.2.3절에서 기술한 바와 같이 식 5.16과 같이 계산되었다.

$$P(w_i) = \frac{n_i}{N} \quad (5.16)$$

여기서  $N$ 은 샘플 데이터의 총 수를 의미하며 식 5.16은 간단히  $w_1$  및  $w_2$ 에 포함된 샘플의 빈도수를 전체 샘플 수로 나누어 확률로 표현한 것이다. 이러한 결과로부터 사후확률  $p(w_i|f)$ 는 식 5.17과 같이 계산되며 각 파라미터의 값은 아래와 같이 추정되었다.

$$p(w_i|f) = \frac{p(f|w_i)P(w_i)}{\sum_{j=1}^2 p(f|w_j)P(w_j)} \quad (5.17)$$

$$p(f|w_1) \sim N(-0.850, 0.146), P(w_1) = 0.514$$

$$p(f|w_2) \sim N(0.898, 0.441), P(w_2) = 0.486$$

Fig. 5.14는 요인들의 분포로부터 추정된 파라미터를 기반으로 하는 두 확률 밀도함수를 도시한 것이다. 또한 요인의 확률밀도함수는 두 가우시안 확률밀도 함수가 결합된 형태로 모델링될 수 있으며(Gaussian mixture model) 이는 식 5.18과 같이 표현될 수 있다.

$$p(f) = \sum_{j=1}^2 p(f|w_j)P(w_j) \quad (5.18)$$

이는 식 5.17의 분모항에 해당되는 요인  $f$ 의 주변확률(Marginal probability)에 해당된다. 베이지안 결정규칙을 적용한 최소오류율 분류를 위해 상대적 청정지역  $w_1$ 과 상대적 오염지역  $w_2$ 을 구분 짓는 최적의 결정경계는 -0.1357로 계산되었다. 예를 들어 특정 미생물에 대한 요인점수를 안다고 하면 요인점수가 경계치보다 클 경우는  $w_2$ 에 포함되었다는 결정을, 작을 경우에는  $w_1$ 에 포함되었다는 결정을 내릴 수 있다. 이러한 결정방법은 미생물데이터의 출현패턴에 대한 클래스-조건부 확률밀도함수를 구하는 과정에 직접적으로 이용되었다.



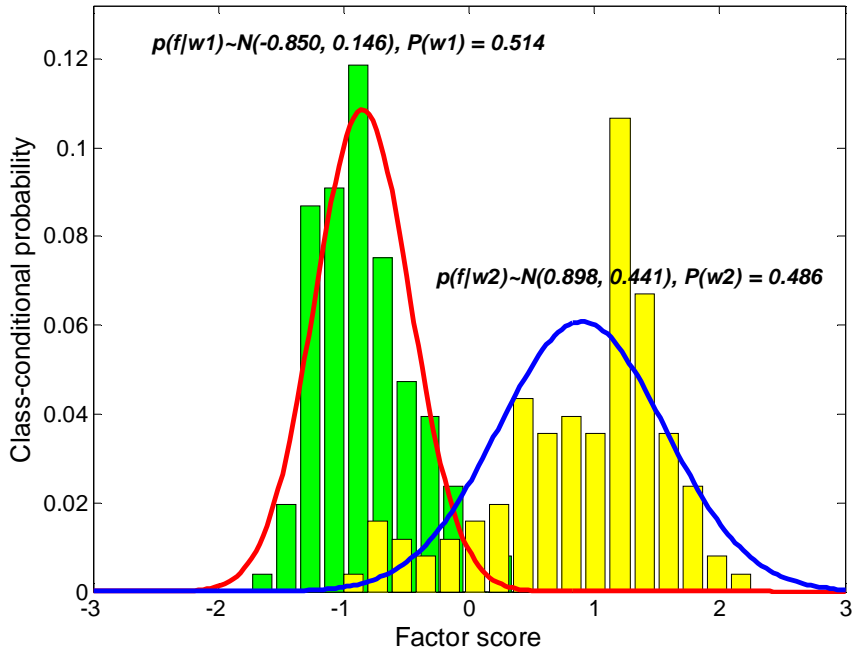


Fig. 5.14. Probabilistic model with the two gaussian distribution functions for the factor scores.

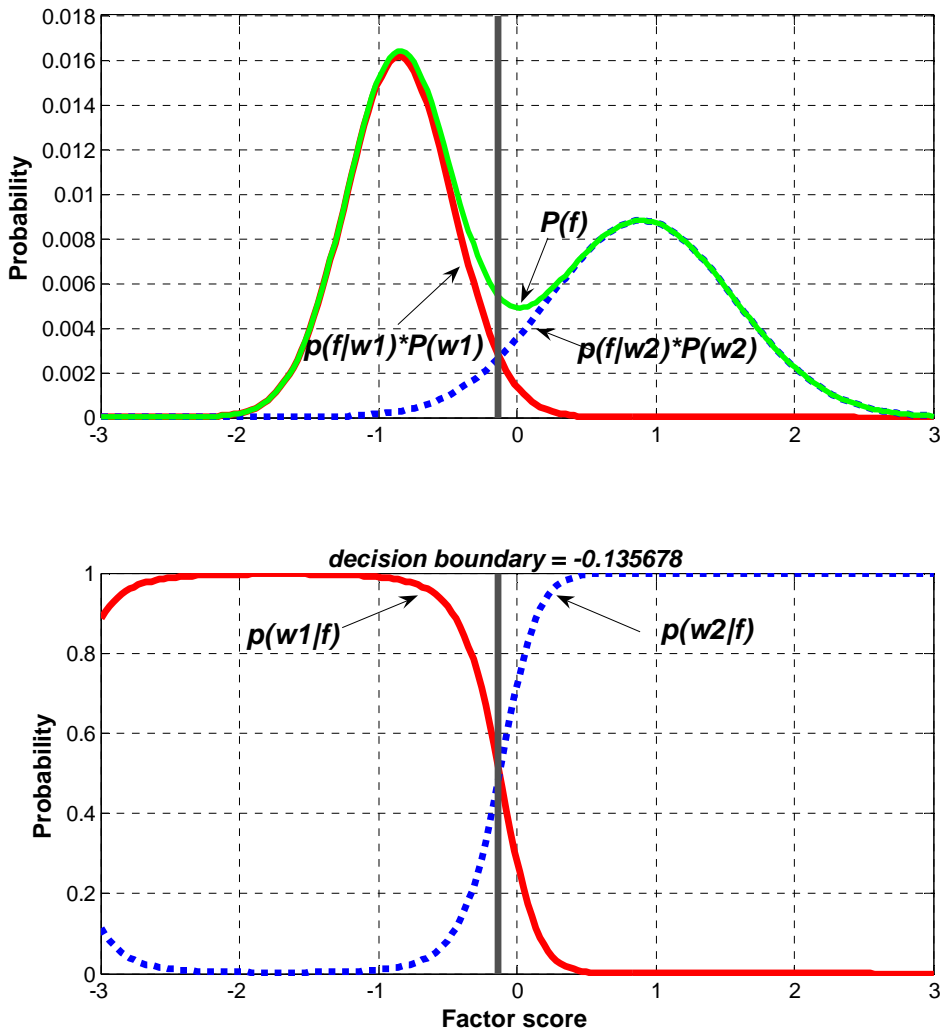


Fig. 5.15. Two-categories classification based on Bayesian decision rule for the factor scores.

### 5.3.4. 확률모형을 이용한 미생물 출현패턴 분석

요인점수의 분포특성을 확률모형을 통한 분석으로 상대적인 오염지역 또는 청정지역을 요인점수의 사후확률을 통해 최적의 결정경계를 얻을 수 있었다. 본 연구에서는 이러한 결정경계를 근거로 하여 각 샘플링 데이터마다 클러스터 정보를 부여하였고 이를 이용하여 미생물의 출현 패턴을 보다 효과적으로 분석하고자 하였다. 각 클러스터 정보는 상대적 청정지역 또는 오염지역의 의미가 부여되었기 때문에 각 군집에서 미생물이 출현할 확률을 나타내는 클래스-조건부 확률은 결국 오염여부에 따른 특정 미생물의 출현확률의 분포를 나타내는 것이므로 미생물의 패턴을 분석하는데 있어 유용한 정보가 될 수 있다.

선행분석의 결과로 요인-미생물 간의 상관성이 높은 미생물은 거의 존재하지 않았기 때문에 미생물 데이터의 수치적 의미보다는 출현빈도수를 분석하는 방법으로 접근하였다. 따라서 5.2.4절에서 논의된 바와 같이 임의의  $k$ 번째 미생물  $M_k$ 에 대해 출현확률만을 고려하여  $w_1$  과  $w_2$ 에 대한 결합 확률인  $P(M_k=1, w_1)$  및  $P(M_k=1, w_2)$ 를 사전확률로 간주하였다. 따라서 오염정도에 따라 출현하는 임의의 미생물  $M_k$ 에 대한 요인점수의 클래스 조건부 확률밀도함수는 각각  $p(f|M_k=1, w_1)$ 과  $p(f|M_k=1, w_2)$ 로 표현하였다. 식 5.24로부터 구한 사후확률  $P(M_k=1, w_i|f)$ 는 요인점수가 주어진 경우  $w_i$ 지역에서  $M_k$ 가 출현할 사후확률을 나타내며 베이저안 결정규칙을 이용하여 미생물의 출현패턴을 4가지로 분류하였다. 분류된 각각의 출현패턴은 미생물이 수질의 오염에 대한 적응성, 내성 등의 결과로 나타난다고 볼 수 있으므로 이를 미생물의 특성으로 간주하였다. 이러한 결과로부터 특성화한 미생물의 출현 패턴을 다음과 같이 정리하였다.

- (1) 상대적 청정지역에서 주로 출현하는 미생물 : P1
- (2) 상대적 오염지역에서 주로 출현하는 미생물 : P2
- (3) 오염정도와 상관없이 출현하는 미생물 : P3
- (4) 오염정도와 상관하여 출현하는 미생물 : P4

115종에 대한 미생물을 특성화한 결과 P1에 해당하는 미생물은 상대적 청정지역에서의 출현빈도수가 오염지역에 비해 높은 특성이 있었다. 반면, P2에 해당하는 미생물은 청정지역에 비해 오염지역에서 높은 출현률을 보였다. P3와

P4로 분류된 미생물의 경우에는 두 지역에서 출현하는 빈도수가 비슷한 경향이 있었다. P4로 분류되는 미생물의 경우에는 이러한 성향과 더불어 요인과 미생물간의 상관성이 비교적 높은 특성을 가지고 있으므로 직접적으로 수질오염의 예측에 활용될 수 있다. 그러나 본 연구에서 분석조사된 115종의 미생물에 대해서 단 1종인 "*Acineotbacter* sp."만이 P4로 분류되었다. 이 미생물과 요인간에는 양의 상관관계( $r=0.43$ ,  $p<0.001$ )를 보였으나 높은 상관성을 가지고 있지는 않았다. Fig. 5.16 ~ Fig. 5.19의 (A)는 상기 기술한 4가지 특성을 가지고 있는 대표적인 미생물에 대한 요인점수의 분포를 나타낸다. 각 패턴에 대한 대표적 미생물은 "*Pandora* sp.", "*beta-proteobacterium*", "*Acinetobacter* sp.", "*Clostridium* sp."로서 각각 P1, P2, P3, P4 패턴의 데이터 분포를 가시적으로 확인하기 위해 활용되었다. Fig. 5.16 ~ Fig. 5.19의 (B)는 미생물의 특성화 과정에서 분석에 이용된 히스토그램 및 두 개의 가우시안 확률모델로 근사된  $w_1$  및  $w_2$ 에 대한 클래스-조건부 확률밀도함수를 나타낸다.

결과에서 보는 바와 같이 P1과 P2는 매우 상반된 경향을 나타내고 있으므로 이러한 미생물의 특성이 환경평가에 충분히 활용될 수 있음을 알 수 있었다. 특히 생물지표의 경우 수질의 오염도가 높을수록 발견되는 종의 수가 저하하여 생물지표를 이용한 환경평가에 한계가 있다. 반면 미생물의 경우 P2의 경우 20여종 이상의 미생물이 검출되었기 때문에 생물지표와 더불어 미생물 지표를 이용한다면 보다 우수한 환경평가가 가능하리라 사료된다.

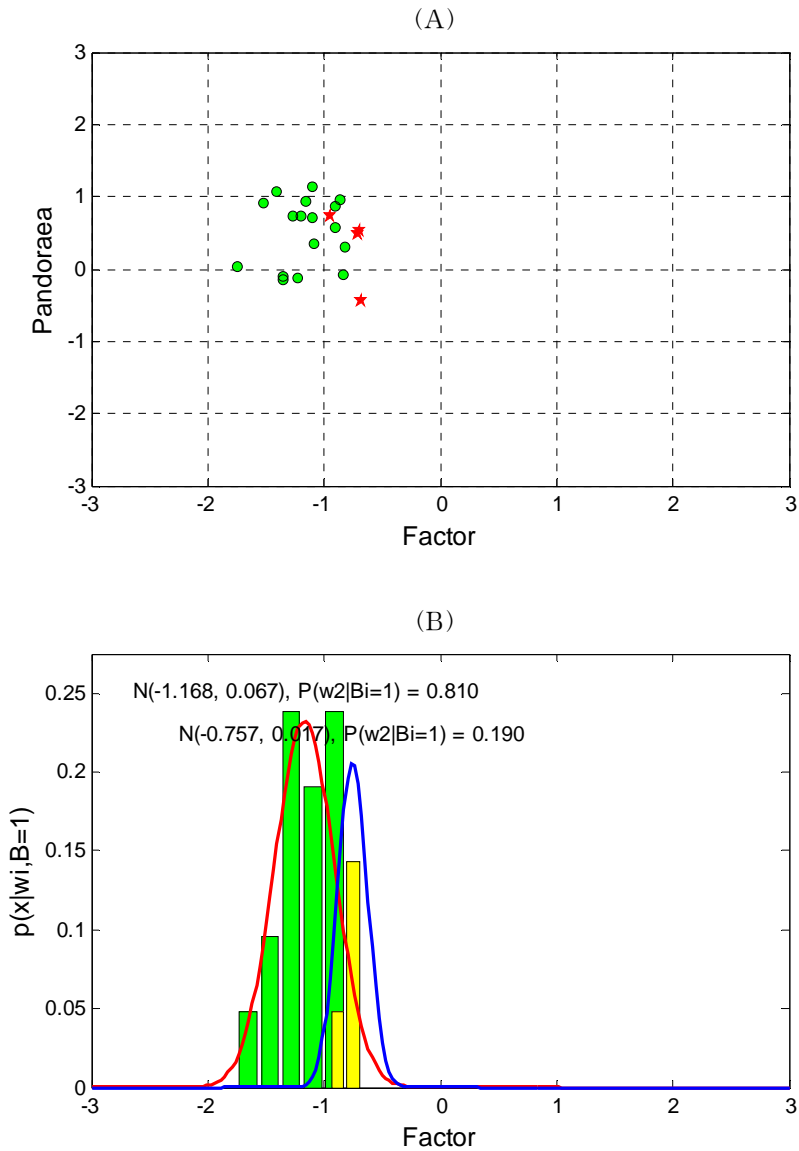


Fig. 5.16. Probability distribution in terms of characterizing microbial community pattern as P1. (A): Joint distribution of two variables; (B): A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.

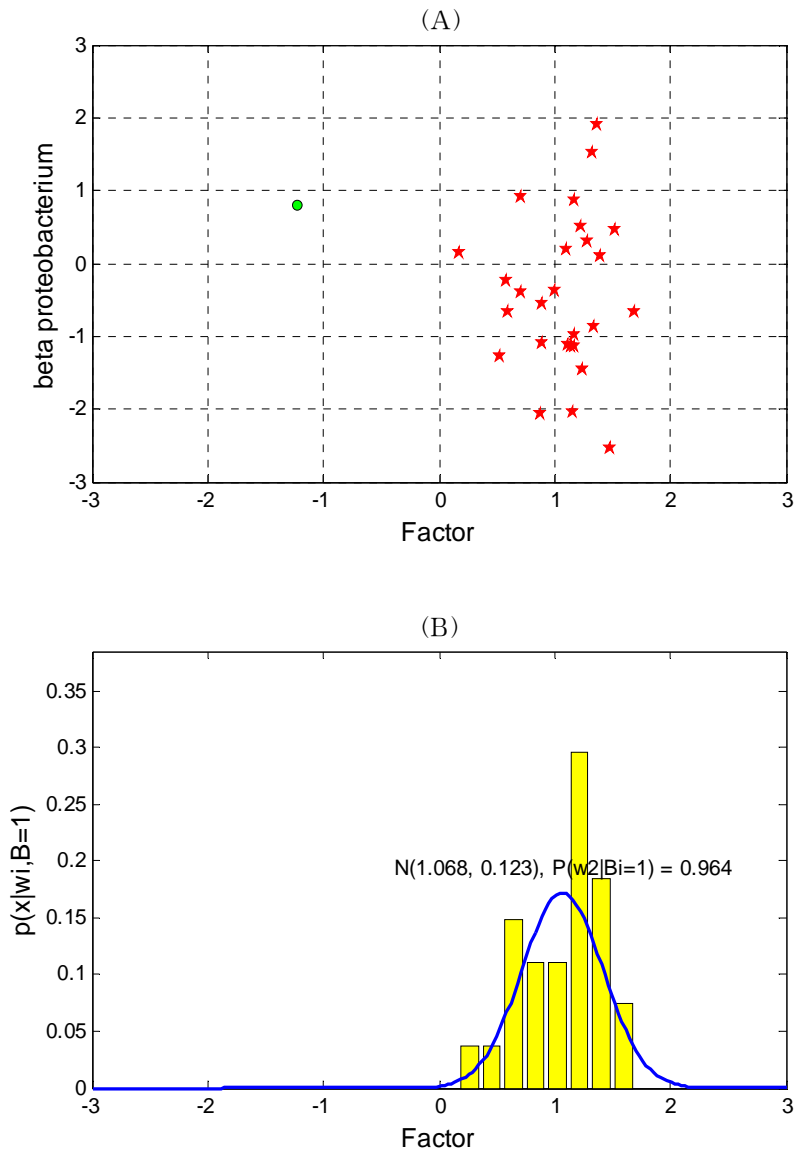


Fig. 5.17. Probability distribution in terms of characterizing microbial community pattern as P2. (A): Joint distribution of two variables; (B): A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.

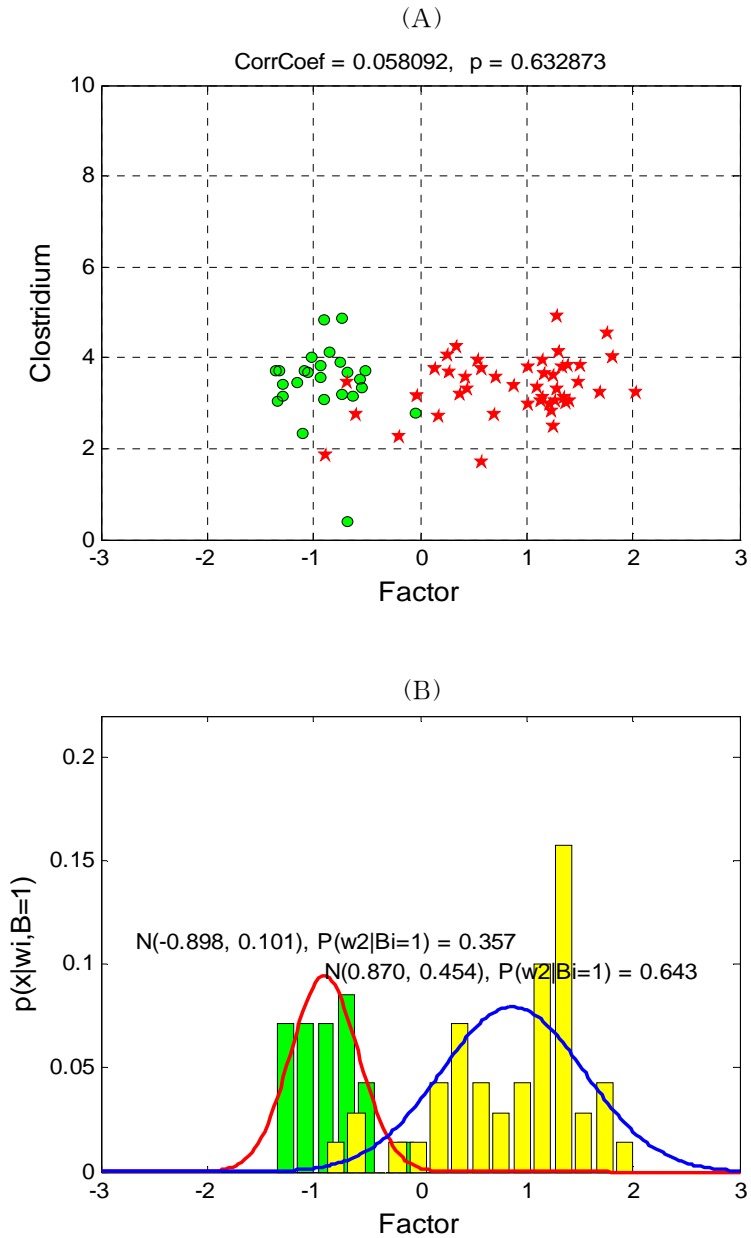


Fig. 5.18. Probability distribution in terms of characterizing microbial community pattern as P3. (A): Joint distribution of two variables; (B): A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.

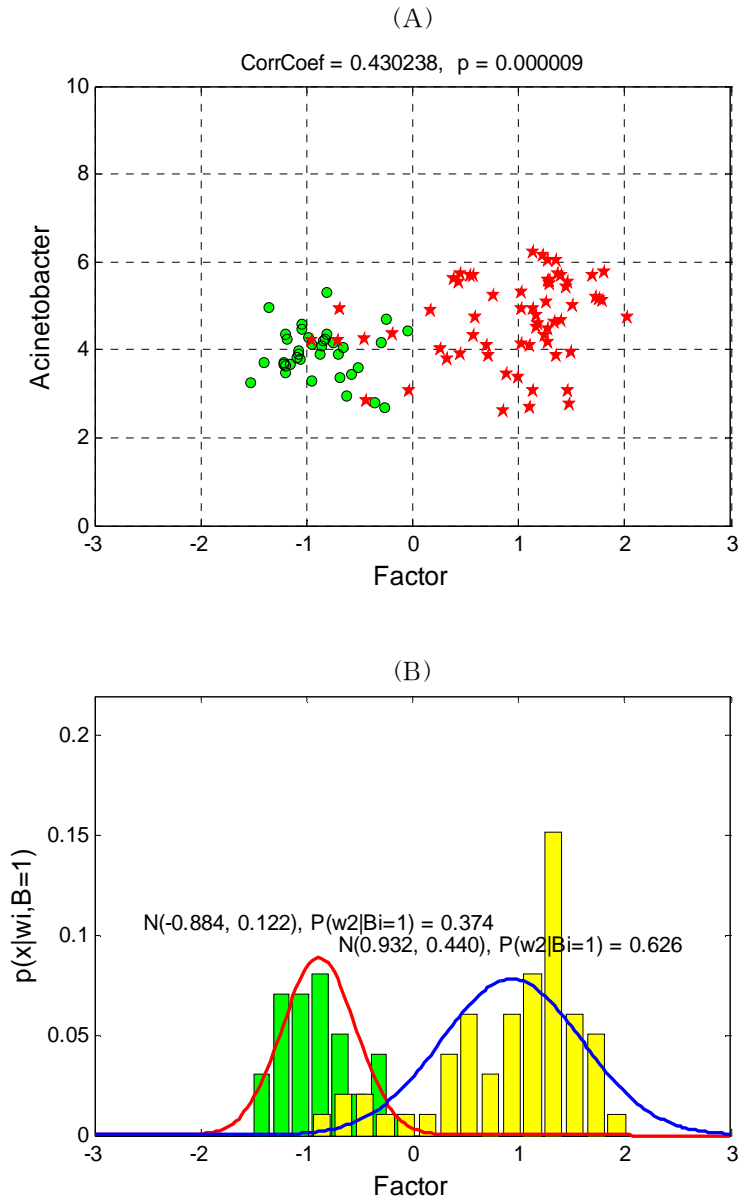


Fig. 5.19. Probability distribution in terms of characterizing microbial community pattern as P4. (A): Joint distribution of two variables; (B): A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.



### 5.3.5. 상대적 중요도 평가를 통한 주요 미생물 추출

5.3.4절에서 미생물을 상대적 수질 오염도에 대한 출현빈도를 이용하여 미생물을 특성화하였다. 이를 통해 얻은 결과로부터 P1과 P2의 특성을 가지는 미생물들은 환경평가의 지표로서 활용가치가 높다고 판단된다. 그러나 미생물의 다양한 종류에 비해 샘플의 수가 적기 때문에 보다 신뢰되는 환경평가의 지표로서의 미생물 군을 선별하기 위해서는 이들의 독립적인 특성과 함께 미생물 각각에 대한 상대적인 관계도 고려되어야 한다. 예를 들어 두 종의 미생물이 서로 상반된 P1과 P2의 패턴을 가지고 있으나 동일한 수질 상태에 대해 서로 같은 판정을 하는 경우나 이와 반대의 경우가 있다. 따라서 이러한 미생물들이 실제 환경평가에 활용되기 위해서 각 미생물의 특성이 잘 반영되어 올바른 환경평가가 가능한 미생물들을 선별하는 것이 필요하였다.

본 연구에서는 5.2.5절에서 기술한 바와 같이 미생물 개개의 특성과 같이 상대적 특성 등을 종합적으로 고려하여 미생물의 상대적인 중요도를 측정할 수 있는 평가시스템을 구성하였으며 이를 통해 지수화가 가능한 미생물들을 선정하였다. 각 변수의 상대적 중요도의 측정에는 각각의 미생물 변수에 대한 특성들이 PNN기반의 분류기를 통해 예측결과에 얼마나 잘 반영이 되는지를 측정하는 것으로 이를 반복적인 모의실험으로부터 평가하였다. Fig. 5.20은 상대적 중요도를 평가한 결과로부터 1차적으로 선별된 미생물과 측정된 상대적 중요도를 도시한 것이다. 총 44종의 미생물에 대하여 각 특성에 대한 구분 없이 높은 상대적 중요도 순으로 정렬하여 나타내었다. 여기서 60% 이상(0.6<sup>2</sup>이상)의 상대적 중요도를 가지는 미생물들을 다시 선별하여 이들을 상대적 청정지역( $w_1$ )을 평가하는데 중요도가 높은 미생물과 상대적 오염지역( $w_2$ )을 평가하는데 중요도가 높은 미생물을 다시 구분하여 Fig. 5.21과 Fig. 5.23에 각각 도시하였다.

상대적 청정지역에서는 Uncultured Verrucomicrobiales, *Stenotrophomonas maltophilia*, *Lactobacillus delbrueckii* 및 *Pandiraea* sp. 등의 순으로 높은 상대적 중요도를 가지는 13종의 미생물이 선정되었다. 선정된 미생물이 발견되었던 환경상태를 알아보기 위해 각 미생물에 대한 요인점수의 분포를 Fig. 5.22와 같이 나타내었다. 여기서 박스좌우 점선의 끝단은 평균치를 기준으로 표준편차의 2.698배에 해당하는 위치를 나타내며 박스의 좌우끝단은 데이터의 25% 경계와 75% 경계를, 박스 가운데의 실선은 데이터의 중앙값을 나타낸다. 미생

물 Uncultured Verrucomicrobiales의 경우 데이터의 50%가 형성되어있는 범위는 -1.15에서 -0.85 근방에 있으며 대부분의 요인점수가 -1 근방에 밀집되어 상대적 청정지역의 대표적 미생물로서 평가될 수 있음을 확인하였다. 다른 미생물 역시 대부분 -1 근처에 데이터가 밀집되어 있어  $w_1$ 지역에서의 상대적 중요도가 높은 순으로 선정된 미생물들이 대체로 각 미생물의 요인점수의 분포가 상대적 청정지역 방향으로 크게 편향되어 있음이 밝혀졌다.

상대적 오염지역에서 60%이상의 중요도를 가지는 미생물 변수들은 *Bdellovibrio* sp.를 비롯하여 *Herbaspirillum* sp., *Duganella* sp. 등의 순으로 총 21종의 미생물이 보고되었다. 각 미생물이 출현할 경우의 환경상태를 간접적으로 확인하기 위해 요인 축 상에서의 출현밀도를 분석한 결과 대부분의 미생물에 대한 요인점수의 중앙값이 1에 근접하고 있음을 알 수 있었다(Fig. 5.24). 상대적으로 오염된 지역에서 중요한 역할을 수행하는 미생물의 분포패턴은 일괄적으로 높은 요인점수를 가지는 방향으로 바이어스 되어있음을 알 수 있으며 *Spingobium* sp.는 비교적 데이터 분포가 넓음에도 불구하고 중앙값이 상대적으로 높은 요인점수를 가지는 방향으로 바이어스 되어 있었다.

60% 미만의 중요도를 가지는 나머지 10종의 미생물 변수는 대부분 P3나 P4에 해당되는 미생물이었다. P3 및 P4에 해당하는 미생물의 특성은 오염지역에 크게 구분 없이 일반적으로 출현하는 미생물들로서 요인점수의 분포는 Fig. 5.25와 같이 분석되었다. 이들의 경우 앞서 살펴본 미생물과는 달리 요인 값이 0인 중앙에 분포하고 있으며 분산도 상대적으로 크게 주어짐을 확인할 수 있었다. 여기서 Uncultured Bacteroidetes의 경우 출현패턴이 P1으로 특성화되었으나 상대적 중요도 평가로부터  $w_1$  또는  $w_2$ 지역의 대표 종으로 선정되지 못했다. 이는 상대적 평가로부터 오염상태의 예측능력이 다른 미생물에 비해 상대적으로 저조했기 때문으로 사료된다. 60% 미만의 중요도를 가지는 10종의 미생물을 대상으로 상관성을 분석해 본 결과 P4에 가장 근접한 특성을 가지는 미생물은 *Acinetobactor* sp. 한 종 뿐이었으며 요인과의 상관성은  $p < 0.001$ 에서의 0.43으로 높게 나타나지는 않았다(Fig. 5.26). P1, P2의 특성을 가지는 미생물은 수질상태를 판정하는 중요한 정보를 제공할 수 있으나 이러한 정보만으로 오염의 정도를 예측하는 것은 쉽지 않다. 그러나 P4의 특성을 가지는 미생물에 대한 정보를 알 수 있다면 이들 미생물의 특성을 조합함으로써 보다 신뢰성 있는 환경평가의 수행이 가능할 것이라 사료된다.

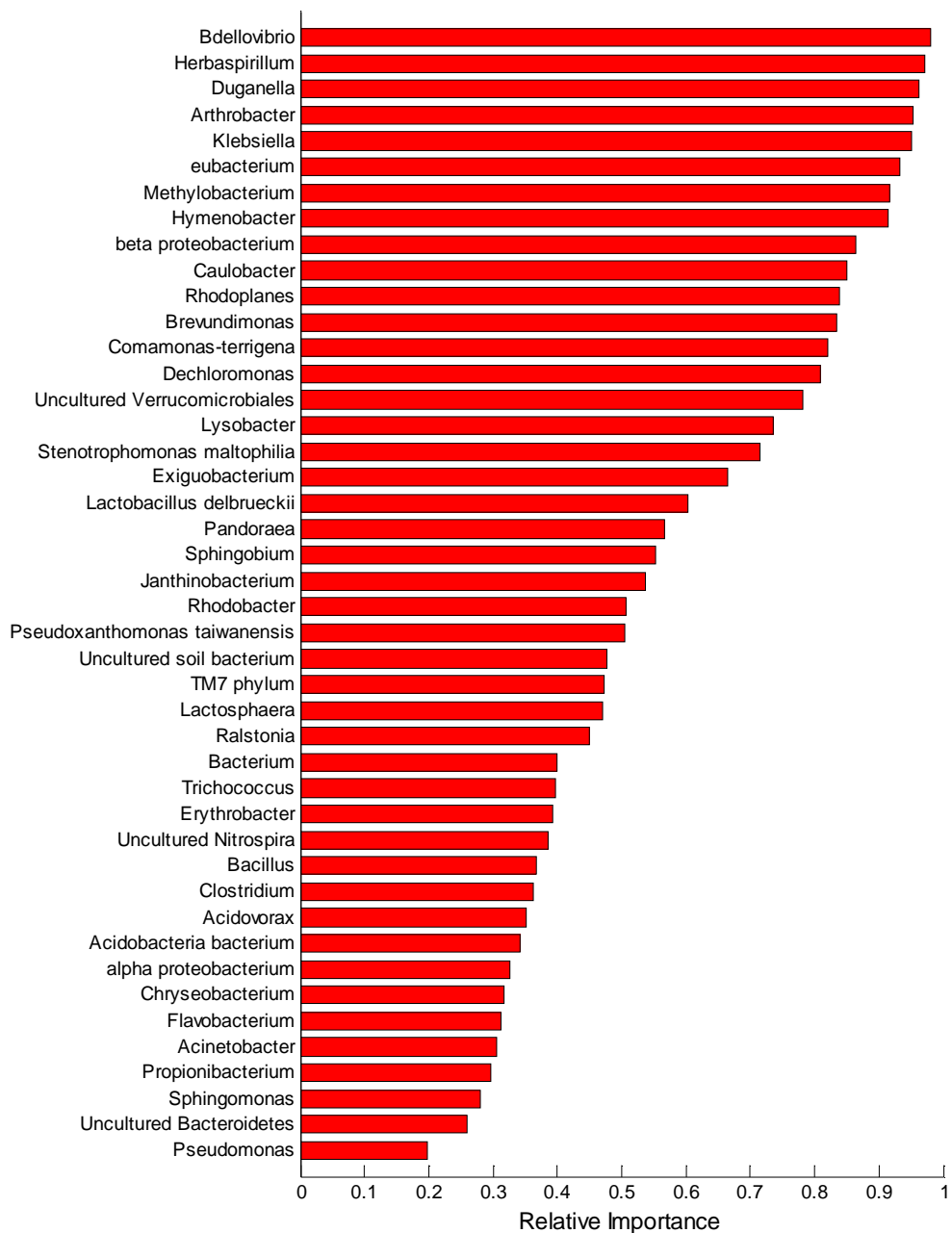


Fig. 5.20. Relative importance of the indicators selected by PNN-based evaluation system.

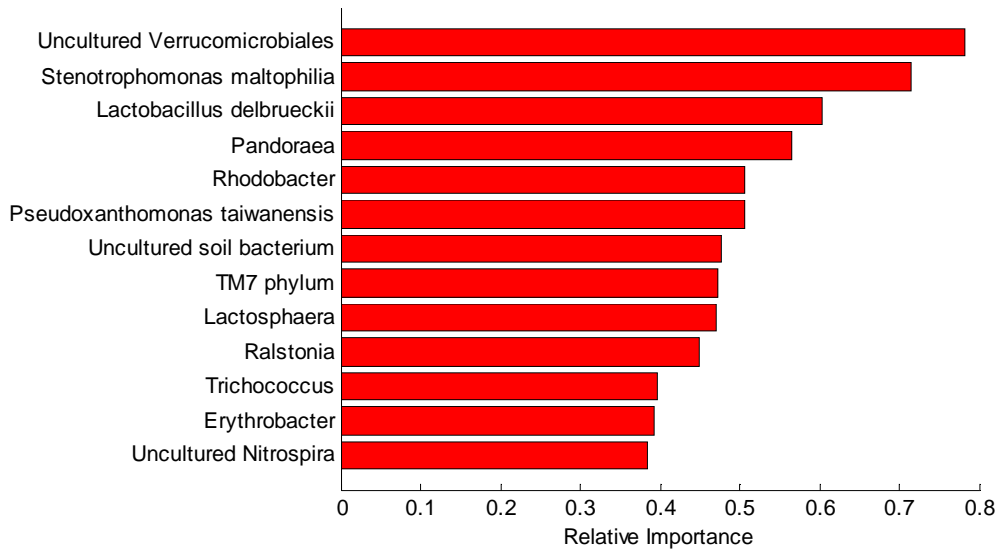


Fig. 5.21. Relative importance of the microbial populations as an indicator in relatively clean site ( $w_1$ ).

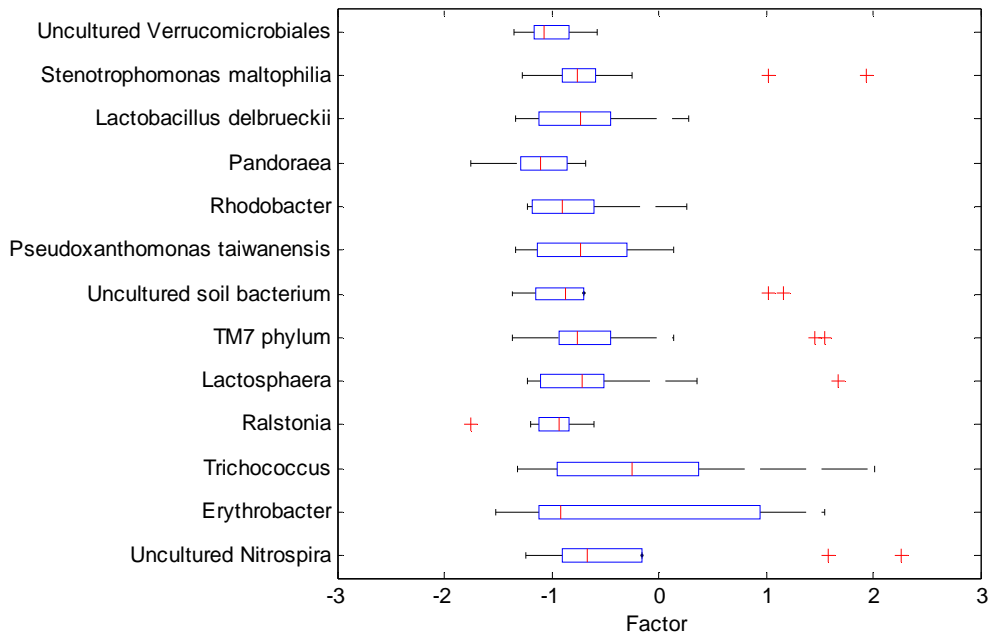


Fig. 5.22. Factor-score distributions for the microbial populations as an indicator in relatively clean site ( $w_1$ ).

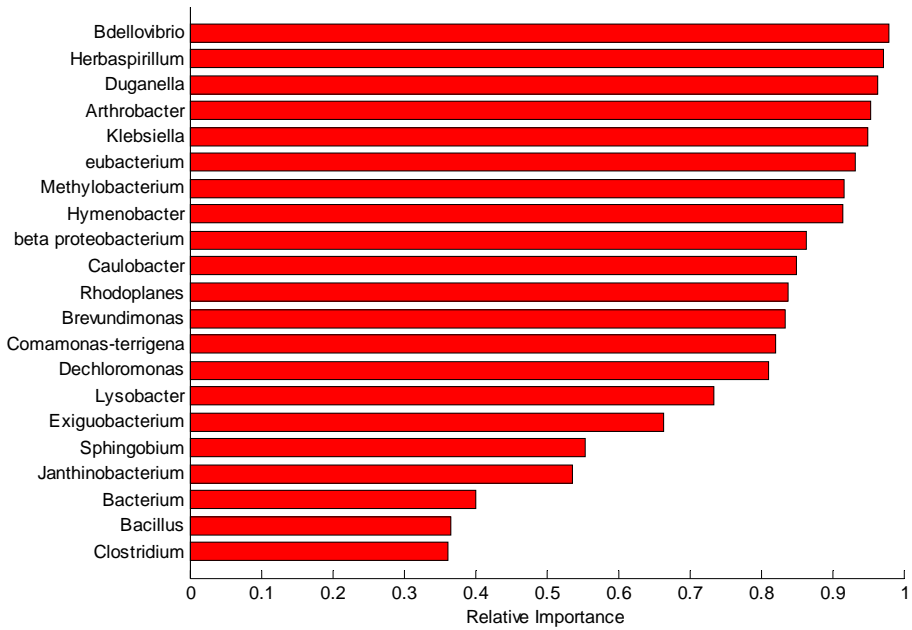


Fig. 5.23. Relative importance of the microbial populations as an indicator in relatively polluted site ( $w_2$ ).

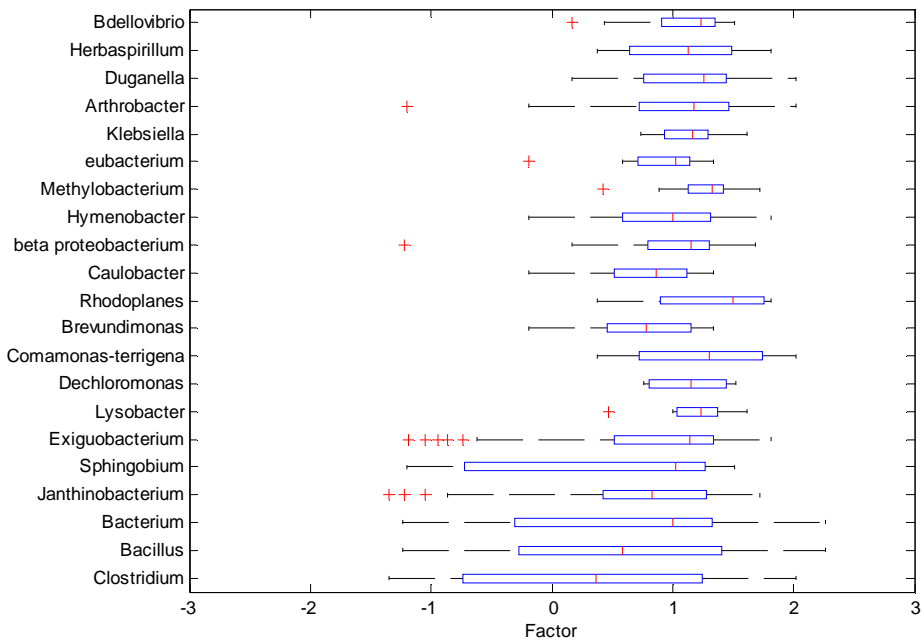


Fig. 5.24. Factor-score distributions for the microbial populations as an indicator in relatively polluted site ( $w_2$ ).

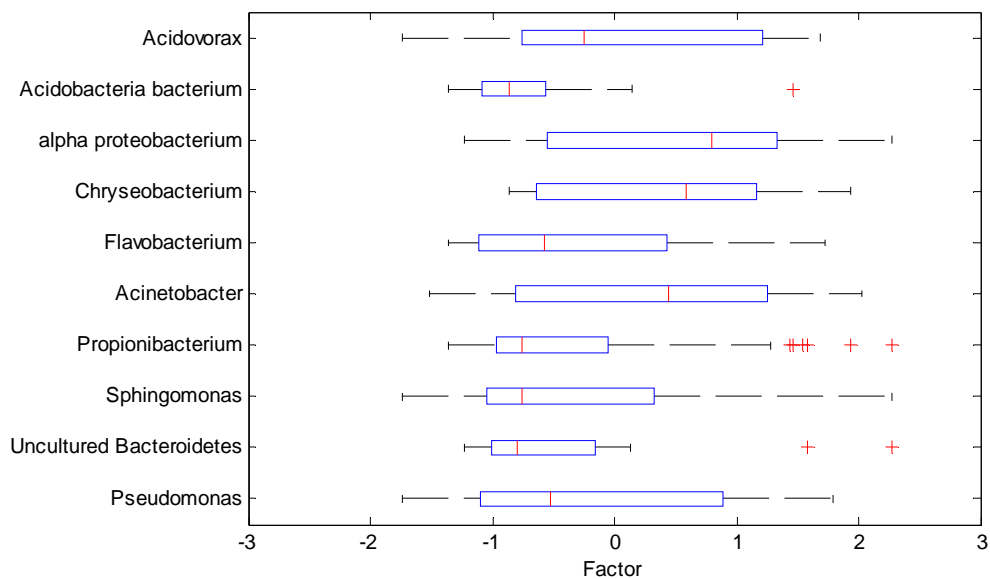


Fig. 5.25. Factor-score distributions for the microbial populations as a indicator in both clean( $w_1$ ) and polluted( $w_2$ ) site.

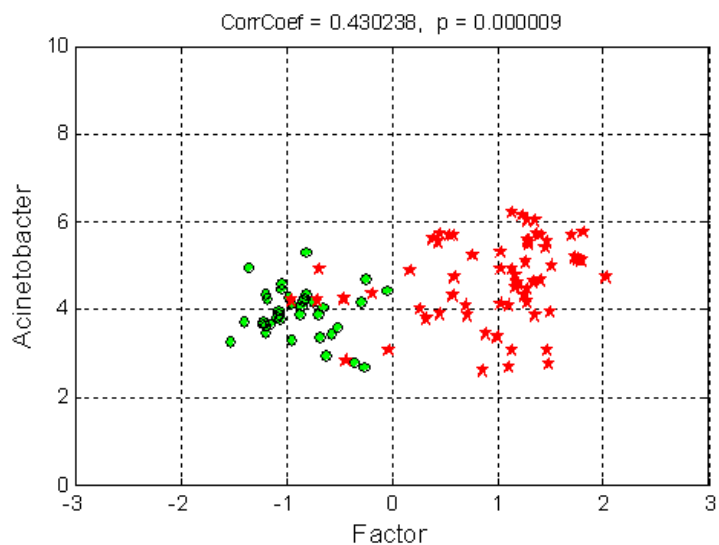


Fig. 5.26. Joint distribution and correlation between the factor and *Acinetobacter* sp. as a predictor in the pattern of P4.

### 5.3.6. 미생물 지수화 및 환경평가

상대적 중요도의 평가로 선정된 미생물은 P1 및 P2 특성을 지니는 미생물로서 수질의 오염의 평가를 위해 다양한 방법으로 활용가능하다. 본 연구에서는 오염정도를 대표하는 지표로서의 미생물들을 열거하고 이들이 지닌 중요도에 따라 지수로 부여하여 최종적으로는 보다 직관적이며 간단한 방법으로 환경평가에 활용하고자 하였다. 이를 위해 우선 선정된 미생물을 오염정도에 따라 10개의 등급으로 분류하기 위해 상대적 중요도로 정렬된 미생물 지표를 요인공간으로 사상시킨 후 이를 다시 0~10사이의 범위를 가지도록 선형 스케일 함수를 이용하여 지수공간으로 사상시켰다. 이때 등급이 높아질수록 낮은 요인점수를 가지도록 하여 높은 등급이 청정도가 증가되도록 하였다. Fig. 5.27은 상대적 중요도에 따라 정렬된 미생물 지표들을 지수공간으로 사상시킨 결과를 보여주고 있다.

여기서 가장 높은 지수점수를 가지는 *Uncultured Verrucomicrobiales*는 상대적 청정지역에서 상대적 중요도가 가장 높은 미생물을 나타내며 가장 낮은 지수점수를 가지는 *Bdellovibrio* sp.는 상대적 오염지역에서 상대적 중요도가 가장 높은 미생물이 된다. 이러한 방법으로 지수차원의 점수를 10개의 등급을 가지도록 이산화하여 선정된 미생물을 분류한 결과로 Table. 5.5와 나타난 바와 같이 미생물 지수화 결과를 얻었다. 환경평가의 지표로 선정된 미생물은 총 21종으로 청정지역의 지표로 9종을 오염지역의 미생물 지표로는 12종으로 주어졌다. 지수화 결과 상대적으로  $I_9$ 과  $I_1$ 에 많은 미생물이 분류되었으며 각각 4종, 5종이 포함되었다. 반면  $I_6$ 에는 미생물이 분류되지 않았는데 이는 청정지역의 대표 미생물이 오염지역에 비해 그 수가 적어 선형 스케일 함수를 적용하는 과정에서 누락된 것으로 사료된다.

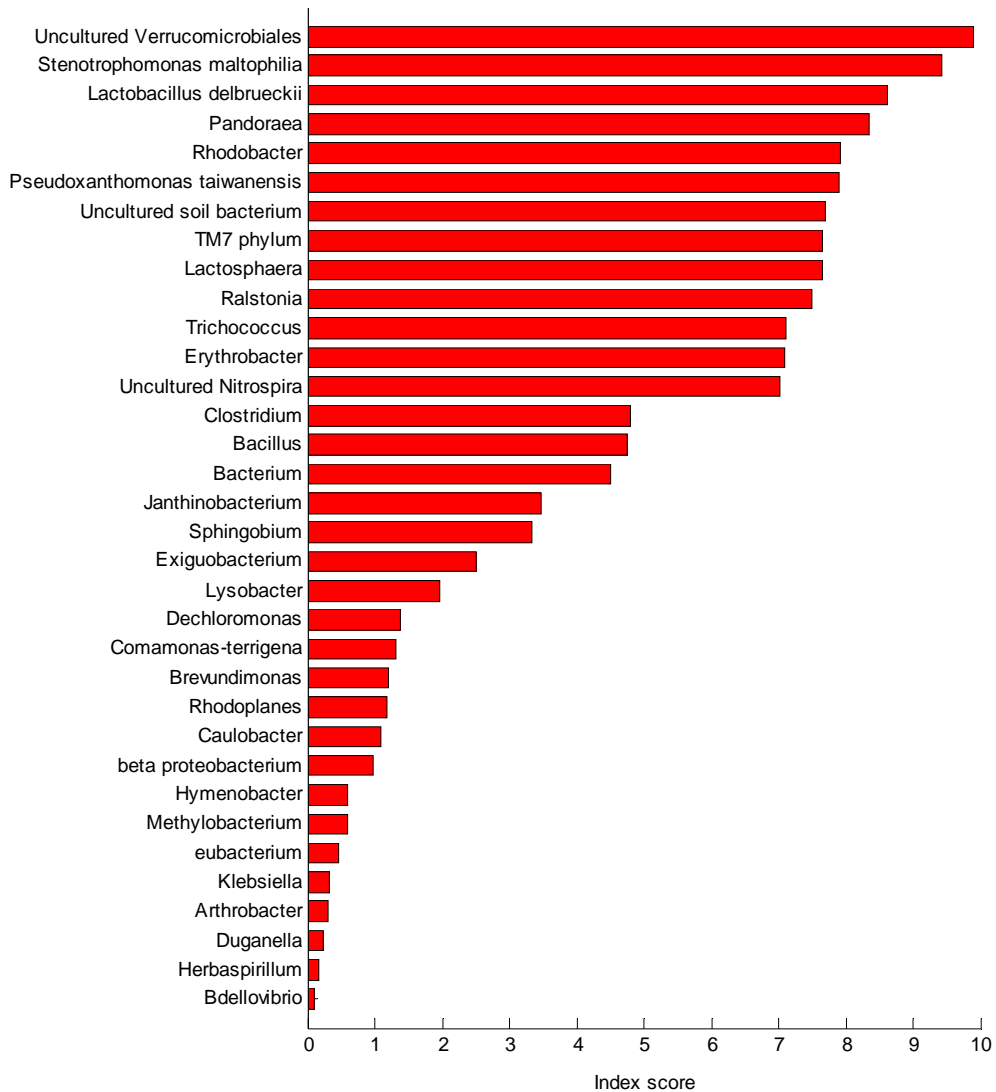


Fig. 5.27. Index scores of the microbial populations selected as an indicator in relatively polluted( $w_2$ ) and clean( $w_1$ ) site as a result of the relative importance evaluation.



Table 5.5. Indices and scores of the selected microbial populations.

| <i>Microbial Indexes</i> | <i>Scores</i> | <i>Microorganisms</i>  |
|--------------------------|---------------|--|
| $I_{10}$                 | 10.0          | Uncultured Verrucomicrobiales<br><i>Stenotrophomonas maltophilia</i>   |
| $I_9$                    | 9.0           | <i>Pandoraea</i> sp.<br><i>Lactobacillus delbrueckii</i>   |
| $I_8$                    | 8.0           | Uncultured soil bacterium<br><i>Pseudoxanthomonas taiwanensis</i><br><i>Ralstonia</i> sp.<br><i>Erythrobacter</i> sp.<br>Uncultured <i>Nitrospira</i><br><i>Ralstonia</i> sp.<br><i>Lactosphaera</i> sp.   |
| $I_7$                    | 7.0           | <i>Trichococcus</i> sp.  |
| $I_6$                    | 6.0           |  |
| $I_5$                    | 5.0           | <i>Bacillus</i> sp.  |
| $I_4$                    | 4.0           | <i>Janthinobacterium</i> sp.<br><i>Sphingobium</i> sp.   |
| $I_3$                    | 3.0           | <i>Exiguobacterium</i> sp.<br><i>Caulobacter</i> sp.   |
| $I_2$                    | 2.0           | <i>Lysobacter</i> sp.<br><i>Dechloromonas</i> sp.<br><i>Comamonas terrigena</i>  |
| $I_1$                    | 1.0           | <i>Rhodoplanes</i> sp.<br><i>Brevundimonas</i> sp.<br>beta proteobacterium<br><i>Caulobacter</i> sp.<br><i>Methylobacterium</i> sp.<br><i>Herbaspirillum</i> sp.<br><i>Arthrobacter</i> sp.<br><i>Klebsiella</i> sp.<br><i>Hymenobacter</i> sp.<br><i>Duganella</i> sp.<br><i>Bdellovibrio</i> sp. |

지수가 부여된 미생물을 이용하면 5.2.6절에서 설명한 바와 같이 관측 데이터로부터 지수가 부여된 미생물의 밀도 값을 이용하여 이들의 가중 합을 계산함으로써 지수점수를 계산할 수 있다. 이때 지수점수는 5대 환경변수들 간에도 밀접한 관련이 있는데 이는 5.3.1절에서 분석한 바와 같이 요인과 각 환경변수들 간의 회귀분석을 통해 얻은 선형 관계식을 이용하여 요인을 매개체로 미생물 지수로부터 각 환경변수들 간의 관계를 알 수 있다. 따라서 미생물 지수를 이용하여 얻은 지수점수를 이용하여 실질적으로 Turbidity, BOD, T-P, T-N 및  $\text{NO}_3^-$ -N의 값이 어느 정도의 범위 내에 있는지를 확인할 수 있다. Table 5.6은 미생물 지수와 각각의 환경변수와의 관계 및 지수점수의 범위를 보여주고 있다.

Table 5.6. Numerical relation between microbial index and each of the environment variables.

| <i>Index</i> | <i>Turbidity</i> | <i>BOD</i>  | <i>T-P</i> | <i>T-N</i>   | <i>NO<sub>3</sub><sup>-</sup>-N</i> |
|--------------|------------------|-------------|------------|--------------|-------------------------------------|
| 10           | 0.00~2.40        | 0.00~2.39   | 0.00~0.11  | 0.00~3.26    | 0.00~5.25                           |
| 9            | 2.41~3.53        | 2.39~3.38   | 0.11~0.30  | 3.26~5.09    | 5.25~7.34                           |
| 8            | 3.53~5.01        | 3.38~4.66   | 0.30~0.52  | 5.09~7.69    | 7.34~10.12                          |
| 7            | 5.01~6.98        | 4.66~6.31   | 0.52~0.78  | 7.69~11.42   | 10.12~13.83                         |
| 6            | 6.98~9.59        | 6.31~8.45   | 0.78~1.09  | 11.42~16.75  | 13.83~18.78                         |
| 5            | 9.59~13.07       | 8.45~11.20  | 1.09~1.44  | 16.75~24.36  | 18.78~25.38                         |
| 4            | 13.07~17.68      | 11.20~14.76 | 1.44~1.86  | 24.36~35.24  | 25.38~34.18                         |
| 3            | 17.68~23.80      | 14.76~19.36 | 1.86~2.35  | 35.24~50.78  | 34.18~45.92                         |
| 2            | 23.80~31.92      | 19.36~25.31 | 2.35~2.92  | 50.78~72.99  | 45.92~61.58                         |
| 1            | 31.92~42.71      | 25.31~32.98 | 2.92~3.59  | 72.99~104.72 | 61.58~82.47                         |

본 연구에서 개발한 미생물 지수의 유효성 검증을 위하여 수집된 데이터를 이용하여 환경평가를 실시하였다. 우선 수집된 데이터 중에서 지수화 된 미생물 데이터만을 선별하였고 식 5.28의 지수계산식을 이용하여 지수점수를 계산하였다. 지수점수의 결과가 지역별 오염도를 얼마나 잘 반영하는지를 알아보기 위해 각 지역별로 평균 지수점수를 구하였고 이를 평균 요인점수와 비교하였다.

Fig. 5.28은 미생물 지수를 이용한 지수점수와 요인점수를 각 지역에 대한 평균치를 얻어 도시한 것이며 여기서 각 샘플 지역은 요인점수를 기준으로 오염차순으로 정렬하여 판독이 용이하도록 하였다. 또한 Table 5.7은 각 지역에서의 지수점수와 요인점수 및 지수점수를 이용하여 얻은 요인점수의 추정 값을 나타낸다. 지수점수와 요인점수를 비교한 결과 중간 범위의 오염도를 보이는 LiK, InJ, HJD, Duk와 오염이 매우 심한 곳으로 추정되는 SSI를 제외하고는 전반적으로 요인점수와 지수점수로부터 추정된 요인점수는 전반적으로 비슷한 결과를 보였다. 중간 범위에서 오차가 많이 나는 것은 지수로 부여된 미생물이 P1과 P2의 특성을 가지는 것으로 이들의 정보로는 중간범위를 잘 설명할 수 없는 것으로 사료된다. 또한 SSI 지역은 출현되었던 미생물이 P2의 특성을 가지는 미생물이 P1의 특성을 가지고 있는 미생물에 비해 상대적으로 적은 종이 출현했기 때문에 오차가 많이 발생된 것으로 사료된다.

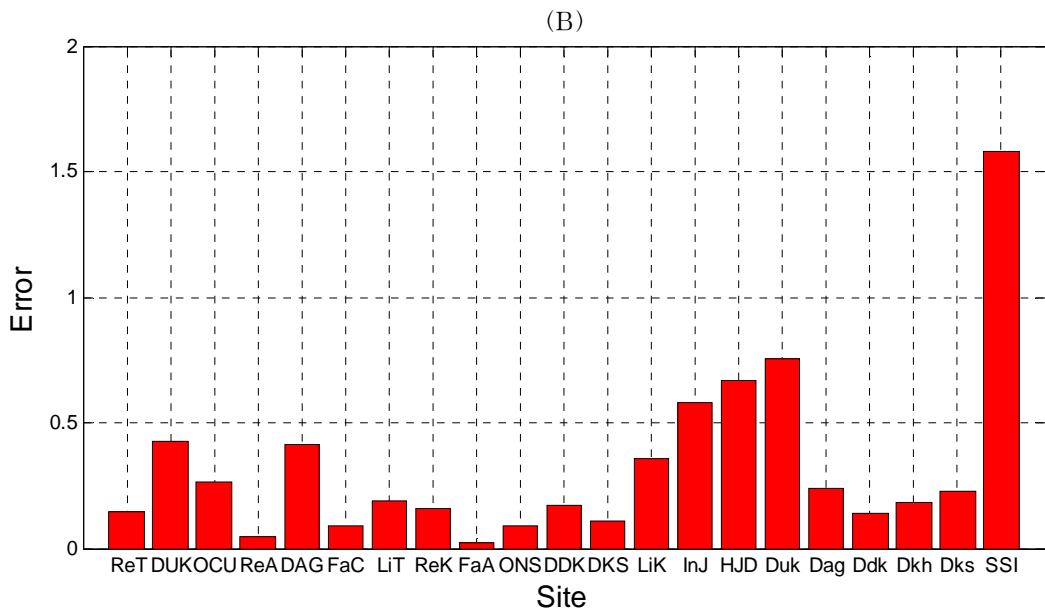
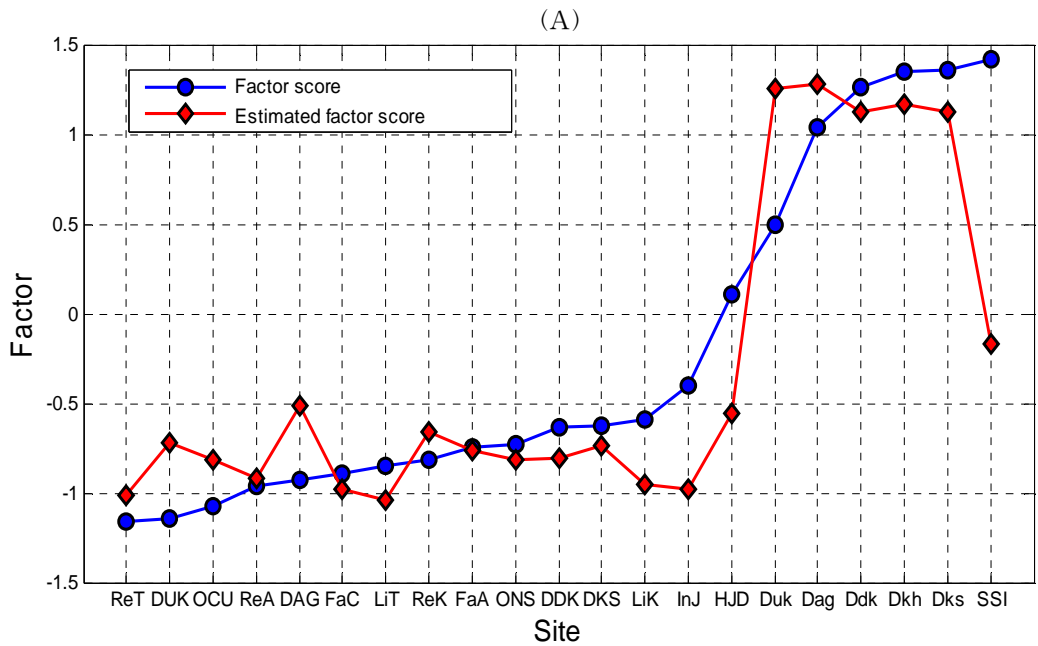


Fig. 5.28. An assessment of water quality using microbial indices. (A) Comparison of average scores between factor and estimated factor evaluated by using microbial index in each sampling site (B) Estimation error of the factor score in each sampling site.

Table 5.7. Comparison of average index scores with average factor scores for the sampling sites.

| <i>Group</i> | <i>Site</i> | <i>Factor score</i> | <i>Estimated factor score</i> | <i>Index score</i> |
|--------------|-------------|---------------------|-------------------------------|--------------------|
| $w_1$        | ReT         | -1.157154           | -1.010701                     | 7.696562           |
|              | DUK         | -1.141357           | -0.768813                     | 7.15781            |
|              | OCU         | -1.076308           | -0.952678                     | 7.567329           |
|              | ReA         | -0.964738           | -0.915362                     | 7.484215           |
|              | DAG         | -0.927401           | -0.726625                     | 7.063846           |
|              | FaC         | -0.89095            | -1.071089                     | 7.831062           |
|              | LiT         | -0.848284           | -1.089596                     | 7.872281           |
|              | ReK         | -0.817104           | -0.684166                     | 6.969279           |
|              | FaA         | -0.742444           | -0.882136                     | 7.410211           |
|              | ONS         | -0.724743           | -0.881765                     | 7.409387           |
|              | DDK         | -0.637922           | -0.944106                     | 7.548235           |
|              | DKS         | -0.625321           | -0.843286                     | 7.323682           |
| $w_2$        | LiK         | -0.592938           | -1.07855                      | 7.847679           |
|              | InJ         | -0.401249           | -1.064167                     | 7.815646           |
|              | HJD         | 0.105592            | -0.771516                     | 7.163831           |
|              | Duk         | 0.496651            | 1.166896                      | 2.84646            |
|              | Dag         | 1.04063             | 1.228643                      | 2.708932           |
|              | Ddk         | 1.260975            | 1.066342                      | 3.07042            |
|              | Dkh         | 1.348111            | 1.083073                      | 3.033155           |
|              | Dks         | 1.355037            | 1.053452                      | 3.099131           |
| SSI          | 1.417262    | -0.230243           | 5.958269                      |                    |

## 5.4. 결론

본 연구에서는 환경평가의 한 방법으로 수질상태에 대한 지표로서의 주요 미생물의 선정방법 및 지수화 방법을 제안하였다. 특히 상관관계의 규명이 쉽지 않은 고차원의 환경변수 및 미생물 변수의 관측데이터를 요인분석 및 확률적 접근을 통해 분석하고 신뢰되는 오염도 평가를 위해 활용가치가 높은 환경 및 미생물 변수들의 추출방법을 제시하였다. 또한 환경평가를 위한 방법으로 분류 및 평가 기법을 이용한 미생물의 지수화 방법 및 수질 평가방법을 제안하였다. 지수화된 미생물에 대하여 주어진 관측데이터를 이용하여 환경평가를 시행한 결과 그 가능성을 충분히 입증하였으며 이에 제안된 방법에 대한 종합적인 결론을 정리하면 다음과 같다.

1. 임의의 미생물 개체수가 특정 환경요인과 밀접한 관련이 있음을 가정하고 상관성을 분석하여 관계식을 얻어 미생물 지수를 개발하고자 하였으나 상관성 분석이 어려웠다. 이는 미생물 변수가 환경변수에 비해 변수의 수가 많으며 missing value가 많기 때문인 것으로 사료된다.
2. 확률분석 기법을 이용하여 산정한 미생물 지수를 이용하여 환경을 평가한 결과 Inj HJD등을 포함한 중간범위를 제외한 범위에서는 대체적으로 좋은 예측결과를 얻을 수 있었다.
3. 지수점수를 산출하기 위해 필요한 미생물의 종류가 청정정도 및 오염정도가 높은 지역과 연관되어 있어, 이러한 미생물들은 그 지역의 환경상태를 잘 설명할 수 있으나 중간범위에서는 그 설명력이 저하되어 이로 인해 중간범위에서의 오차범위가 크게 나타난다고 사료된다.
4. 보다 효율적인 수질예측 지수를 얻기 위해서는 오염지역의 분류 그룹을 증가시켜 중간 범위의 환경을 대변하는 미생물 종들에 대한 지수를 확보함으로써 보완할 수 있을 것으로 사료된다.

## 제 6장 대형무척추동물 지수 및 환경평가를 통한 유효성 검토

본 연구에서는 환경요인과 생물요인을 종합하여 수리분석 및 모델링을 통한 효율적인 통합수질지수를 개발하고자 한다. 이를 위해서는 각 분류군의 군집특성을 우선적으로 파악하는 것이 중요하다. 제 5장에서 미생물 군집분포를 특성화하여 수질평가를 위한 미생물 지수를 제안하였다. 제안된 방법은 미생물 데이터가 수집된 지역의 수질상태를 예측성능을 확인함으로써 환경평가를 위한 충분한 가치가 있음을 검증하였다. 이에 더불어 제 6장에서는 제안된 미생물 지수화 방법을 보다 일반화시키기 위해 저서성 대형무척추동물에 적용하여 제안된 방법의 유효성을 검토하고자 하였다. 이를 통해 제안된 지수화 방법을 다른 분류군에 확장·적용함으로써 각 분류군의 수질상태에 대한 특성을 잘 반영하는 통합지수를 개발하고 나아가 생산자-소비자-분해자의 공존관계를 반영하는 종합적인 수질평가의 가능성을 확인해 보고자 하였다.

미생물 지수를 이용한 환경평가의 유효성을 검증하기 위해 이용된 대형무척추동물 데이터는 미생물 데이터를 수집하였던 동일한 조건의 담수에서 수집된 85개의 대형무척추동물 군으로 구성되어 있다. 대형무척추동물 데이터는 미생물 데이터에 비해 결측값이 상대적으로 적었으나 환경데이터에 비해서는 여전히 많은 결측값을 포함하고 있어 변수간의 상관성에 기초한 일반적인 통계학적 방법으로는 신뢰할 만한 결과를 얻는 것이 쉽지 않았다. 따라서 미생물의 특성 및 상대적 중요도에 기초한 미생물 지수화 방법을 대형무척추동물에 동일하게 적용하면 수질오염과 관련된 주요 지표생물을 선별할 수 있을 것이며 이들에게 부여한 지수를 이용하여 환경평가에 활용될 수 있을 것이라 사료된다.

한편 대형무척추동물을 이용한 수질평가 연구는 서식처의 담수 환경을 반영하는 종의 출현과 밀도를 이용하여 평가된다. 대형 무척추동물을 이용한 환경평가의 한 방법으로 이용되는 BMWP(Biological Monitoring Working Party)지수는 담수생태계의 환경을 평가하는 지표로서, 환경변화에 민감하고 분류군에 따라 비교적 뚜렷한 내성범위를 가지고 있는 대형 무척추동물의 특성을 잘 반영한 것으로 보고되고 있다. BMWP지수가 부여되는 저서성 대형 무척추동물

균은 수질상태에 따른 지표로서의 역할을 하며 본 논문에서 제안된 방법 역시 이러한 대표 균 들을 선정하여 지수를 부여한다. 따라서 BMWP지수를 이용한 환경평가방법과 제안된 방법으로부터 선정된 대형무척추동물의 지수를 이용한 BMWP환경평가 방법을 비교하면 제안된 방법이 수질상태를 평가하는데 있어 그 타당성이 증명될 수 있을 것이다. 이에 본 논문에서는 제 5장의 방법을 대형무척추동물에 동일하게 적용하여 지수를 부여한 뒤 이로부터 수질평가를 수행하여 얻은 결과를 토대로 통합지수의 가능성을 논의하고자 하였다.



## 6.1. 실험방법

### 6.1.1. 대형무척추동물 지수화

대형무척추동물 데이터는 미생물 데이터를 수집하였던 동일한 조건의 담수에서 수집된 85과의 대형무척추동물군으로 구성되어 있다. 대형무척추동물 데이터는 미생물 데이터에 비해 결측값이 상대적으로 적었으나 환경데이터에 비해서는 여전히 많은 결측값을 포함하고 있다. 대형무척추동물의 지수화는 제 5장의 방법과 동일한 과정을 통해 수행되었다. 5대 환경변수에 대한 요인분석을 통해 추출된 공통요인은 대형무척추동물이 출현하는 수질상태를 설명하는 대표변수로 이용되었다. 또한 대형무척추동물의 출현패턴을 분석하기 위해 필요한 상대적으로 청정지역( $w_1$ )과 상대적으로 오염지역( $w_2$ )에 대한 군집정보도 그대로 이용되었다. 이러한 정보를 토대로 대형무척추동물의 지수화를 위한 실험방법은 다음과 같다.

먼저 공통요인에 대하여 각각 두 군집에 대한 정보를 토대로 대형무척추동물의 출현 패턴을 분류하여 각 대형무척추동물을 특성화하였다. 출현패턴은 두 군집에 대한 대형무척추동물의 관측데이터를 최우추정법을 통해 추정된 가우시안 파라미터를 이용하여 두 개의 가우시안 분포를 가지는 클래스-조건부 확률 밀도함수로 근사하였다. 이러한 정보를 근거로 대형무척추동물에 대하여 베이저안 결정규칙에 따라 각각의 개체를 4개의 특성으로 분류하였다.

다음단계로 PNN기반의 상대적 중요도 평가시스템을 통해 대형무척추동물에 대한 수질상태의 예측능력이 상대적으로 우수한 군들을 선정하였다. PNN은 분류기로서 대형무척추동물의 관측데이터의 80%를 무작위 추출하여 학습을 시킨 후, 학습이 완료된 분류기는 수질상태를 평가하는 예측기로서의 역할을 하는데 관측데이터의 나머지 20%가 대형무척추동물의 상대적 중요도를 평가를 위해 이용되었다. 학습과 평가의 반복적인 모의실험이 총 200회가 수행되는 과정에서 각 변수에 대한 수질상태의 예측성능이 측정되고, 모의실험이 완료되면 이 값은 대형무척추동물 개개의 상대적 중요도로써 주어지는데 상대적 중요도가 높은 대형무척추동물은 P1 및 P2의 특성을 가지는 종으로 분석되었다.

마지막 단계는 선정된 대형무척추동물을 상대적 청정지역 및 오염지역에 대

하여 각각 상대적 중요도에 따라 지수를 부여하는 것으로 우선 상대적 중요도에 따라 정렬된 변수들이 수질평가가 가능하도록 요인공간으로 사상시켰다. 또한 이를 선형 스케일링 함수를 통해 1에서 10의 범위를 가지는 지수공간으로 변환하였고 10분할함으로써 무척추동물의 지수를 부여하였다.

### 6.1.2. 지수의 유효성 검토

대형무척추동물의 지수화 방법의 타당성을 검토하기 위해 본 연구에서는 두 가지의 수질 평가방법을 이용하였다. 첫째, 대형무척추동물의 샘플을 채취한 지역의 수질상태를 나타내는 공통요인에 대하여 각 지역별로 요인점수의 추정 능력을 통해 지수의 타당성을 평가하였다. 이를 위해 각 샘플지역의 관측데이터를 다시 지역별로 분류한 후 이들의 요인점수를 대형무척추동물의 지수로부터 계산된 요인점수의 추정치와 비교하였다.

둘째, 수질평가의 한 방법으로 이용되는 BMWP를 이용하여 평가결과의 유사성을 확인하는 방법으로 지수화 방법의 유효성을 검토하였다. BMWP는 대형무척추동물을 이용한 환경평가의 한 방법으로 현재 수질평가에 활발히 이용되고 있으며 담수 생태계의 환경을 평가하는 지표로서 유용한 것으로 보고되고 있다. 따라서 대형무척추동물의 BMWP를 이용한 평가결과와 제안된 방법을 통해 얻은 지수를 이용하여 환경을 평가한 결과와 비교분석하는 것은 제안된 방법의 유효성을 평가함에 있어 보다 객관적이며 의미 있는 일이라 여겨진다. 이를 위해 21개의 시료 채취 지점별로 BMWP에 제시된 대형무척추동물의 각 family에 대한 내성지수(tolerance score)를 부여한 후 이 내성지수의 합을 구하여 BMWP 점수를 구하였다. 본 연구에서 제시된 대형무척추동물 지수들은 같은 방법을 적용하여 각 family에 부여하여 지수의 합을 구하여 이들의 유사성을 검토하였다.

## 6.2. 결과 및 고찰

### 6.2.1. 확률모델을 이용한 대형무척추동물 출현패턴 분석

수질상태에 따른 대형무척추동물의 출현패턴을 분석하기 위해 상대적인 청정지역  $w_1$  및 오염지역  $w_2$ 으로 분류된 두 클래스에 대하여 대형무척추동물의 출현빈도를 이용하여 클래스-조건부 확률분포를 분석하였다. 이를 위해 두 클래스에 대한 조건하에서 출현하는 무척추동물의 요인점수의 분포를 가우시안 분포를 가지는 확률밀도함수로 근사하였다. 가우시안 파라미터인 평균, 분산 및 사전확률은 식 5.14~5.16의 최우추정법을 통해 추정하였다. 이를 통해 근사된 두 클래스-조건부 확률밀도함수와 각 클래스의 사전확률은 식 5.5의 사후확률을 구하는데 이용되었다. 두 클래스의 사후확률로부터 5.24절의 베이시안 결정규칙을 적용하여 무척추동물의 특성을 다음의 4가지의 패턴으로 분류하였다.

- (1) 상대적 청정지역에서 주로 출현하는 대형무척추동물 : P1
- (2) 상대적 오염지역에서 주로 출현하는 대형무척추동물 : P2
- (3) 오염정도와 상관없이 출현하는 대형무척추동물 : P3
- (4) 오염정도와 상관하여 출현하는 대형무척추동물 : P4

미생물의 경우와 마찬가지로 P1과 P2로 특성화된 대형 무척추동물은 각각 청정수와 오염수의 조건을 가지는 특정 환경에만 서식하는 생물로 간주할 수 있다. 또한 P3과 P4의 특성은 오염에 대한 내성이 강한 생물이라 볼 수 있다. 조사기간 동안 출현한 대형무척추동물 85 family에 대해 특성화하였다. Fig. 6.1 ~ Fig. 6.4는 각각의 패턴별로 분류한 대형무척추동물의 확률적 분석결과를 나타낸 것으로 도표 (A)는 각 대형무척추동물에 대한 요인점수의 분포를 나타내었고, 도표 (B)는 각 대형무척추동물 특성화 과정에서 분석에 이용된 히스토그램 및 두 개의 가우시안 확률모델로 근사된  $w_1$  및  $w_2$ 에 대한 클래스-조건부 확률밀도함수를 나타내었다. 각 패턴에 대해 대표적인 대형무척추동물은 Oligochaeta, Ceratopogonidae, Chironomidae 및 Physidae로서 각각 P1, P2, P3 및 P4 패턴의 데이터 분포를 가시적 나타내었다.

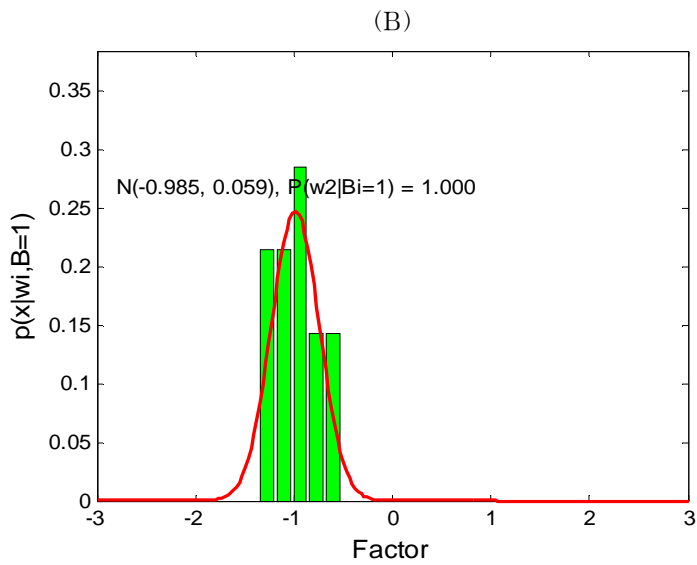
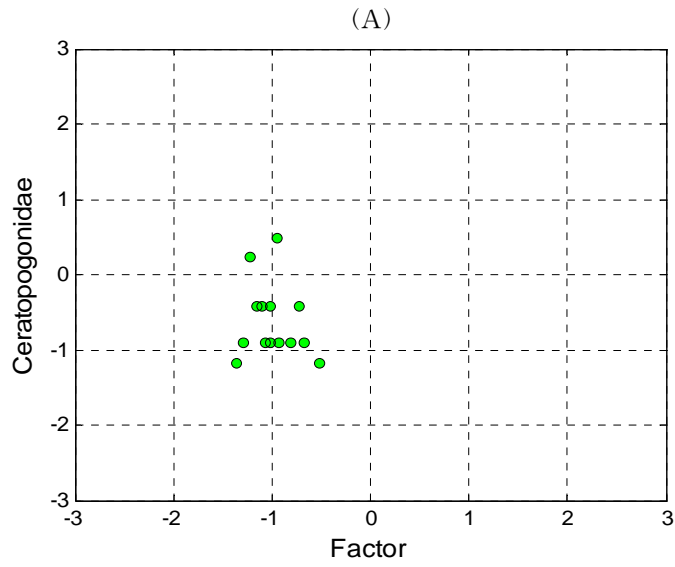


Fig. 6.1. Probability distribution in terms of characterizing macroinvertebrate community pattern as P1. (A) Joint distribution of two variables; (B) A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.

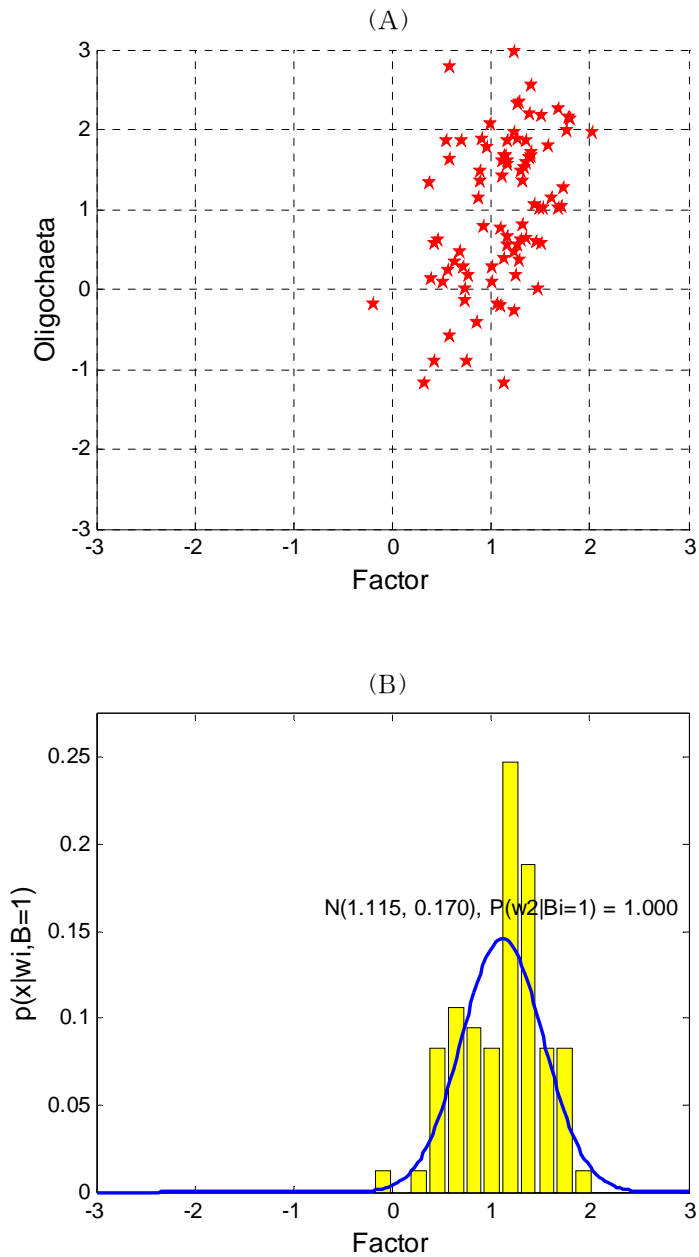


Fig. 6.2. Probability distribution in terms of characterizing macroinvertebrate community pattern as P2. (A) Joint distribution of two variables; (B) A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.

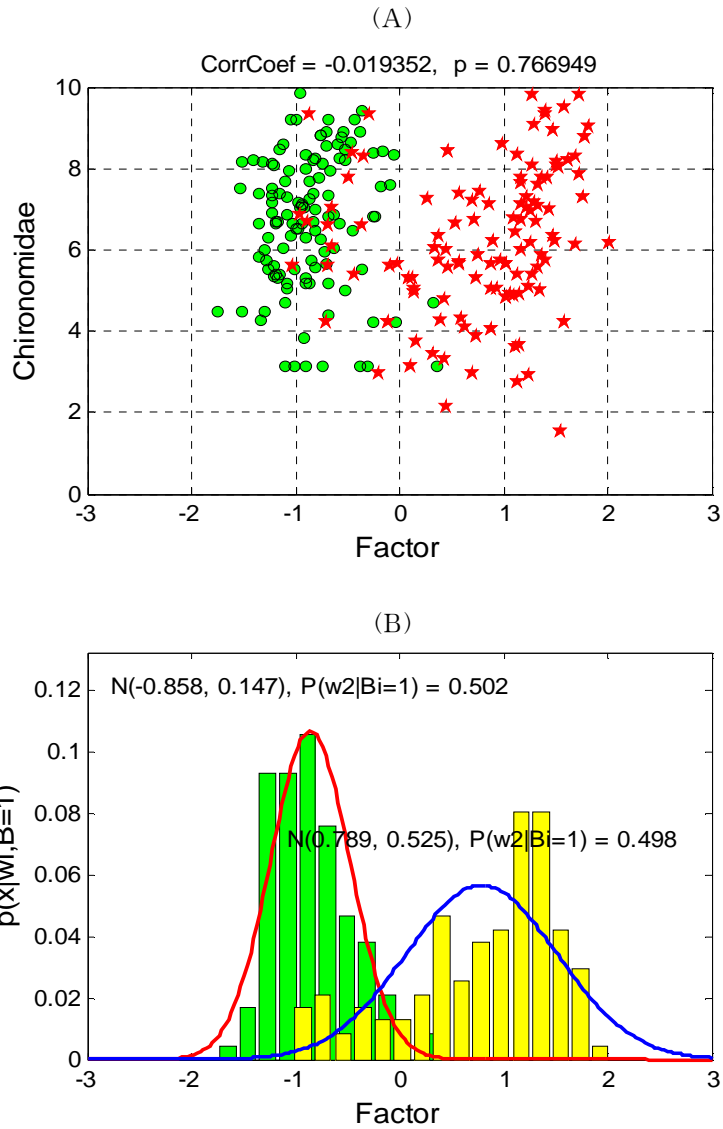


Fig. 6.3. Probability distribution in terms of characterizing macroinvertebrate community pattern as P3. (A) Joint distribution of two variables; (B) A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.

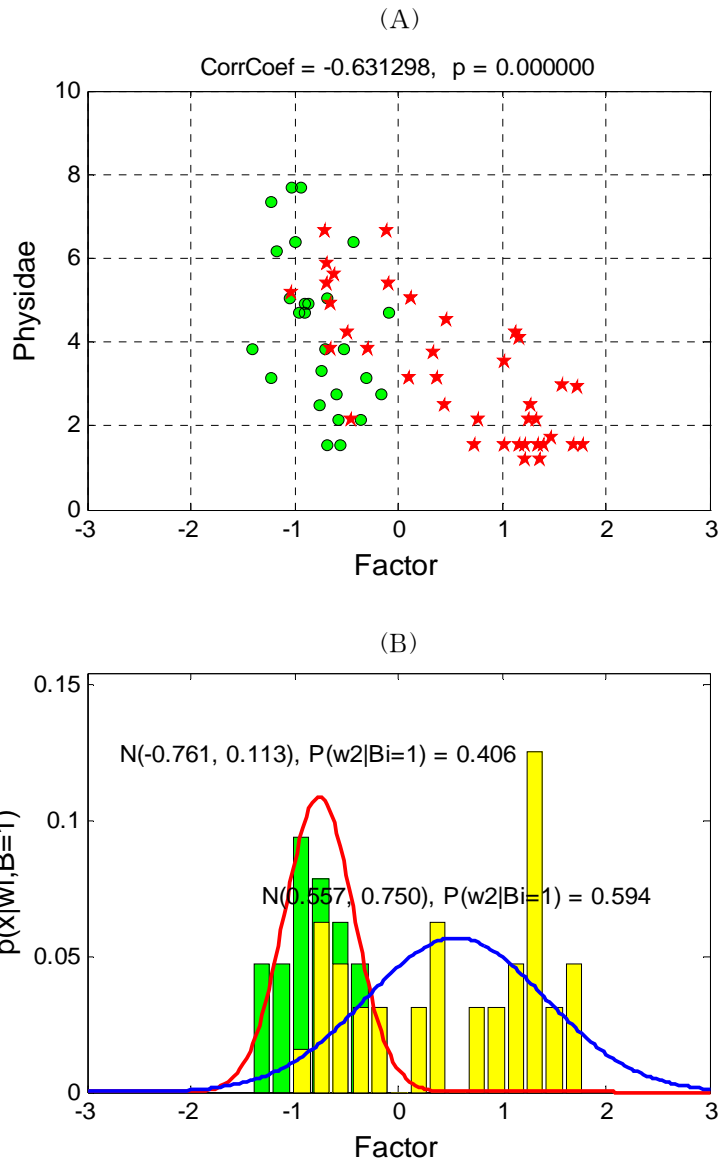


Fig. 6.4. Probability distribution in terms of characterizing macroinvertebrate community pattern as P4. (A) Joint distribution of two variables; (B) A probabilistic model of the joint distribution approximating to gaussian curve through maximum likelihood estimation.

## 6.2.2. 상대적 중요도 평가를 통한 주요 대형무척추동물 추출

대형무척추동물 변수의 상대적 중요도 평가를 수행한 결과 Fig. 6.5와 같은 결과를 얻을 수 있었다. 총 변수의 수는 85이었고 1차적으로 선별된 48개의 변수들에 대한 상대적 중요도는 Fig. 6.5와 같다. 상대적 중요도는 Oligochaeta 가 84.2%로 가장 높게 측정되었다. 60%이상의 중요도를 가지는 대형무척추동물 중 상대적 청정지역( $w_1$ )을 평가하는데 중요도가 높은 대형무척추동물 과 상대적 오염지역( $w_2$ )을 평가하는데 중요도가 높은 대형무척추동물을 다시 구분하여 Fig. 6.6과 Fig. 6.8.에 각각 도시하였다.

상대적 청정지역에서 60%이상의 중요도를 가지는 대형무척추동물은 Chloroperlidae, Nemouridae 및 Ceratopogonidae 등의 순으로 총 18과의 대형무척추동물이 선정되었다(Fig. 6.6). 상대적 중요도평가로부터 선정된 대형무척추동물은 제 5장에서 분석된 미생물과 같이 출현패턴별로 분류가 가능하였는데 여기서 분류된 대형무척추동물은 P1의 특성이 부여된 것으로 조사되었다. 이는 상대적 청정지역에 출현하는 대형무척추동물의 요인점수 분포를 나타내는 Fig. 6.7을 통해 가지적으로 알 수 있다. 이들의 요인 축 상에서의 출현분포를 분석한 결과 대부분의 대형무척추동물의 중앙값이 -1에 근접하고 있는 것으로 나타났다.

상대적 오염지역에서 60%이상의 중요도를 가지는 대형무척추동물은 총 7과로 Oligochaeta, Sarcophagidae, Hirudinidae, Leutridae, Glossiphoniidae, Planorbidae 및 Psychodidae가 선정되었다(Fig. 6.8). 선정된 대형 무척추동물은 모두 P2의 특성을 가지는 것으로 조사되었으며 상대적으로 P1의 특성을 가지는 무척추동물에 비해 상대적으로 적은 family가 선정되었다. 상대적 오염지역에 출현하는 대형무척추동물의 분포 특성 및 패턴을 Fig. 6.9에 나타내었다. 각 대형무척추동물은 요인 축 상에서의 출현밀도를 분석한 결과 모든 대형무척추동물의 중앙값이 1에 근접하고 있는 것으로 나타났다. 결과적으로, 상대적 오염지역에서 중요한 역할을 수행하는 대형무척추동물의 분포패턴은 일괄적으로 높은 오염도를 가지는 방향으로 분포되어 있음을 알 수 있다.



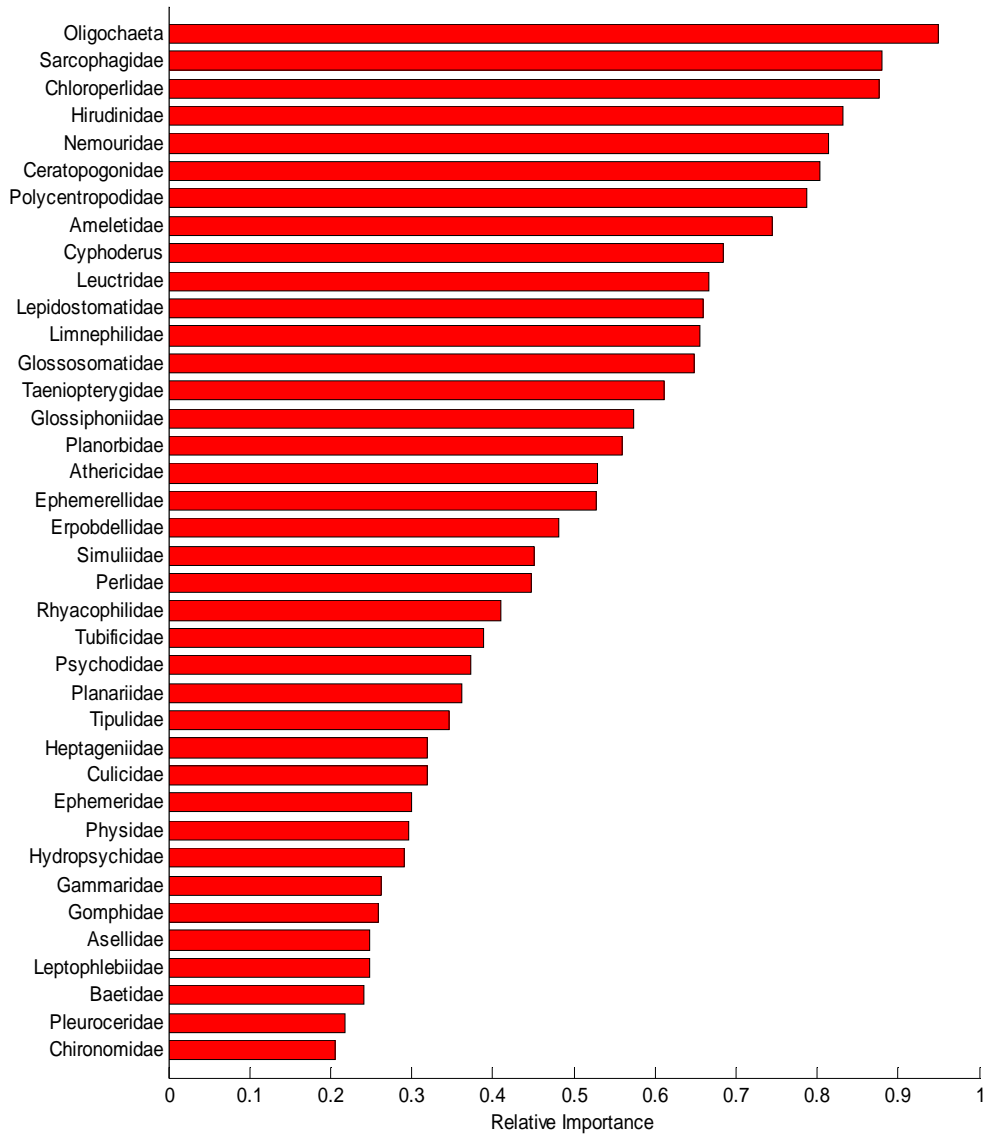


Fig. 6.5. Relative importance of the indicators selected by PNN-based evaluation system.

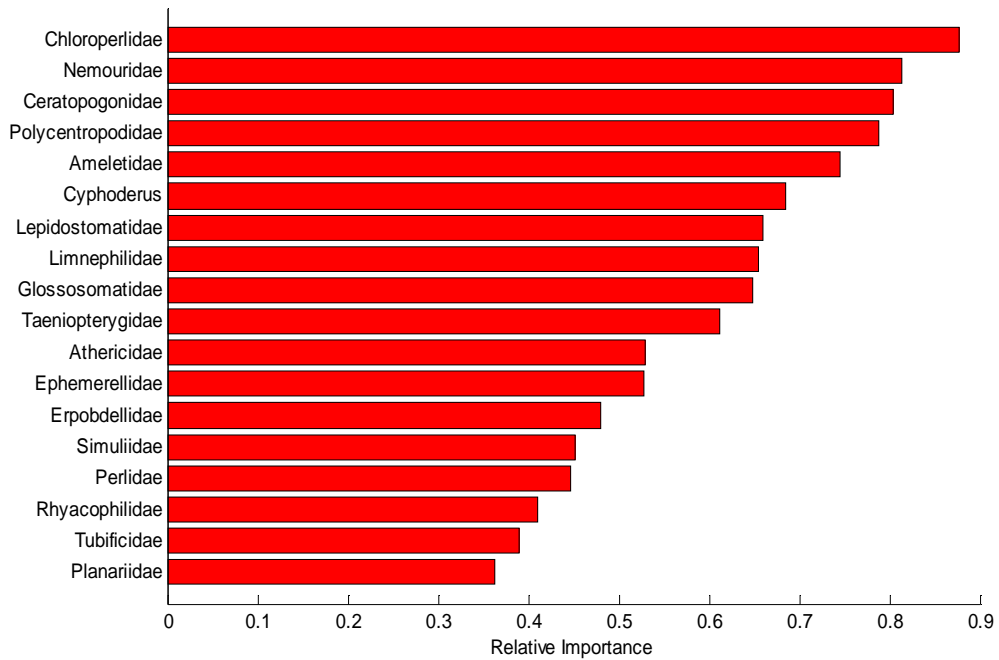


Fig. 6.6. Relative importance of the macroinvertebrate populations as an indicator in relatively clean site ( $w_1$ ).

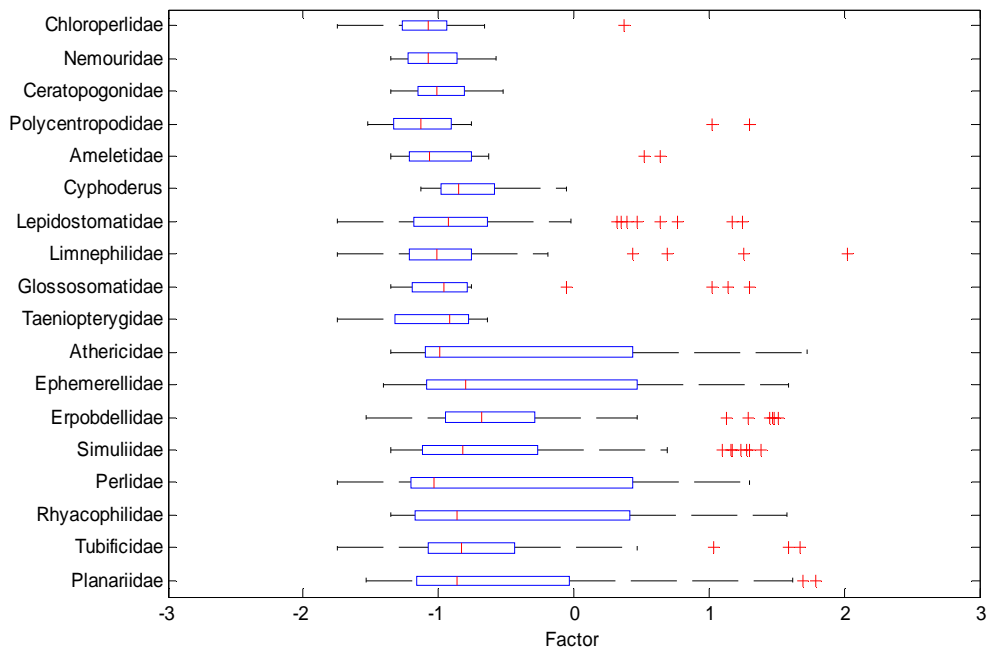


Fig. 6.7. Factor-score distributions for the macroinvertebrate populations as an indicator in relatively clean site ( $w_1$ ).

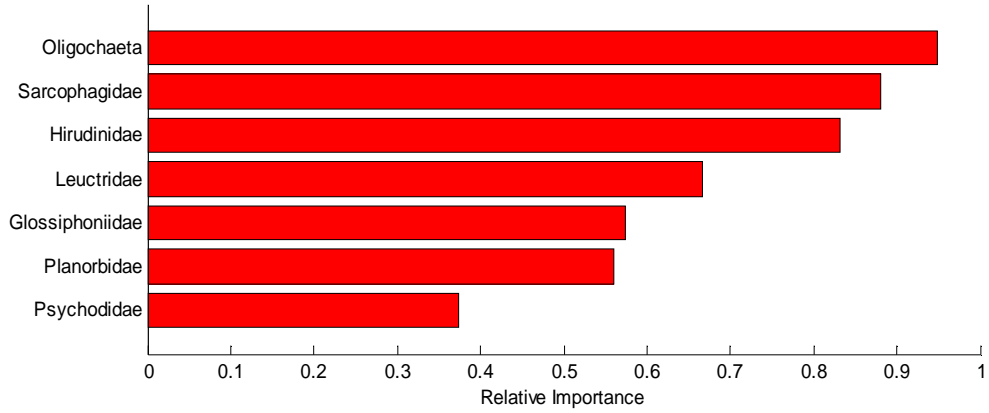


Fig. 6.8. Relative importance of the macroinvertebrate populations as an indicator in relatively polluted site ( $w_2$ ).

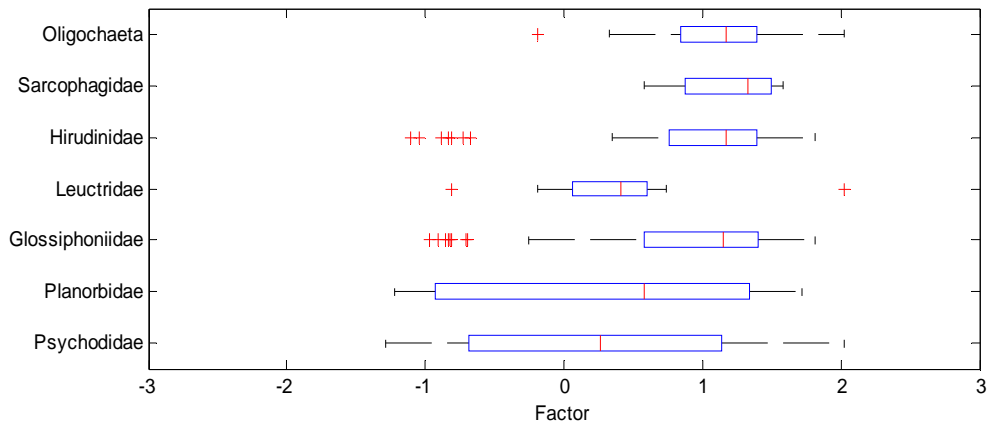


Fig. 6.9. Factor-score distributions for the macroinvertebrate populations as an indicator in relatively polluted site ( $w_2$ ).

이러한 결과로부터 미생물의 경우에는 상대적 오염지역에서 출현하는 미생물 종의 수가 청정지역에서 출현하는 미생물 종의 수에 비해 상대적으로 많은 반면 대형무척추동물의 경우에는 상대적 오염지역에서 출현하는 family가 상대적 청정지역에서 출현하는 family에 비해 상대적으로 적다는 것을 알 수 있다. 이는 대형무척추동물이 미생물에 비해 수질의 상태에 따라 생존에 필요한 조건에 대하여 보다 민감한 반응을 보이기 때문이라 사료된다. 이러한 결과로부터 수질오염이 심한 지역에서 지표로 활용될 수 있는 대형무척추동물의 family가 서식조건이 열악하여 다양하게 보고되지 않을 수 있으며 이로 인해 수질오염이 심한 지역의 환경평가를 위한 충분한 정보를 확보하는 것이 어려울 것이라 여겨진다. 반면 본 연구에서 다루는 미생물의 경우 상대적으로 오염지역에서도 다양한 종이 출현하여 지수를 부여받았으므로 대형 무척추동물에 비해 보다 많은 정보를 제공할 수 있으리라 사료된다.

60% 미만의 중요도를 가지는 13과의 대형무척추동물들의 분포 특성 및 패턴을 Fig. 10에 나타내었다. 대부분 P3나 P4의 분포형태를 가지는 것으로 오염지역에 관계없이 분포하였다 요인 축 상에서의 출현밀도를 분석한 결과 대부분의 대형무척추동물의 중앙값이 0에 근접하여 분포하고 있으며, 분산도 상대적으로 크게 나타났다.

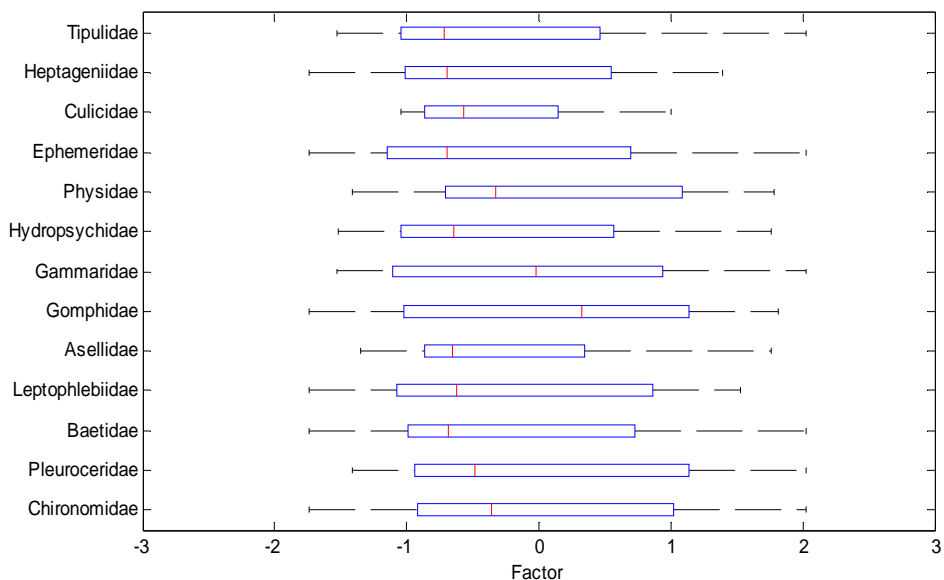


Fig. 6.10. Factor-score distributions for the macroinvertebrate populations as a indicator in both clean( $w_1$ ) and polluted( $w_2$ ) site.

### 6.2.3. 대형무척추동물 지수화 및 환경평가

미생물 지수 선정 방법과 동일한 방법으로 대형무척추동물의 특성이 PNN분류기의 예측성능에 상대적으로 잘 반영되는 26과의 대형 무척추동물을 선정하여 개개의 상대적 중요도를 지수공간으로 사상하여 Fig. 6.11과 같은 결과를 얻었다. 여기서 지수점수 5점 이하의 무척추 동물이 5점 이상의 무척추 동물에 비해 다양성 측면에서 차이가 있음을 알 수 있었다. 무척추동물의 지수는 지수공간을 10개의 등급으로 분할함으로써 오염도에 따라 Table 6.1과 같이 분류하였다. 지수화 결과 6등급 및 4등급에 소속되는 대형무척추동물은 나타나지 않았으며 7등급 이상이 18과로 3등급 이하가 6과에 비해 상대적으로 낮은 지수를 부여받은 대형무척추동물에 비해 적은 종류가 선정되었다.

미생물의 지수화 방법의 유효성은 두 가지의 수질 평가방법을 이용하여 평가되었으며 첫 번째 방법은 대형무척추동물의 샘플을 채취한 지역의 수질상태를 나타내는 공통요인에 대하여 각 지역별로 요인점수의 추정 능력을 통해 평가하였다. 이를 위해 우선 선정된 대형무척추동물에 대한 개별적인 지수와 관측 데이터를 각 시료채취 지점에 따라 분류한 후 식 5.10을 적용하여 지수점수를 산출하였다. Table 2는 각 지역별 요인점수, 대형무척추동물 지수점수 및 이를 요인점수로 환산한 값을 각각 나타내며 Fig. 6.12는 각 지역별 요인점수와 지수로부터 추정된 요인점수를 비교하였다. 여기서 (A)는 각 지역별 요인점수를 비교한 것이고 (B)는 추정오차의 절대치를 나타낸 것이다. 이를 통해 대형무척추동물의 지수를 이용한 수질평가 성능이 미생물 지수에 의한 수질평가 성능에 비해 우수한 것으로 조사되었다. 특히 DUK, DAG 등의 상대적 청정지역에 대해서 미생물의 경우 예측오차가 다소 큰 경향이 있었으나 무척추동물의 경우에는 보다 안정된 예측성능을 보여주었다. 오염지역에서 무척추동물의 지수를 이용한 평가결과 전반적으로 미생물에 비해 상대적으로 많은 오차를 수반하였다. 또한 HJD, Dag 등의 중간 오염지역에서는 미생물의 경우와 마찬가지로 많은 오차를 포함하였으며 특히 SSI지역에서는 미생물의 경우와 같이 매우 큰 오차를 보였다.

결과적으로 대형무척추동물의 지수를 이용한 수질평가는 미생물의 지수를 이용한 수질평가에 비해 상대적 청정지역에서 우수한 성능을 보였으며 오염이 심한 지역에서는 저조한 성능을 보였다. 이는 앞에서 언급한 바와 같이 상대적

오염지역을 잘 설명할 수 있는 대형무척추동물이 다양하게 존재하지 않았기 때문에 사료되며 반면 상대적 청정지역을 대표하는 대형무척추동물은 미생물에 비해 충분히 수질상태를 잘 반영할 수 있는 종류가 발견되었기 때문에 사료된다. 또한 중간지역에서 오차가 많이 발생하는 이유는 지수화된 대형무척추동물 역시 P1과 P2의 특성만을 가지고 있어 이러한 정보만으로 중간지역 수질상태의 설명력이 부족한 것으로 판단된다. 이는 대형무척추동물의 분포특성에 의한 것으로 오염이 심하거나 빈영양 상태의 청정지역에서는 소비자인 대형무척추동물의 종풍부도 및 다양성이 낮은 것에서 기인한 것으로 사료된다.

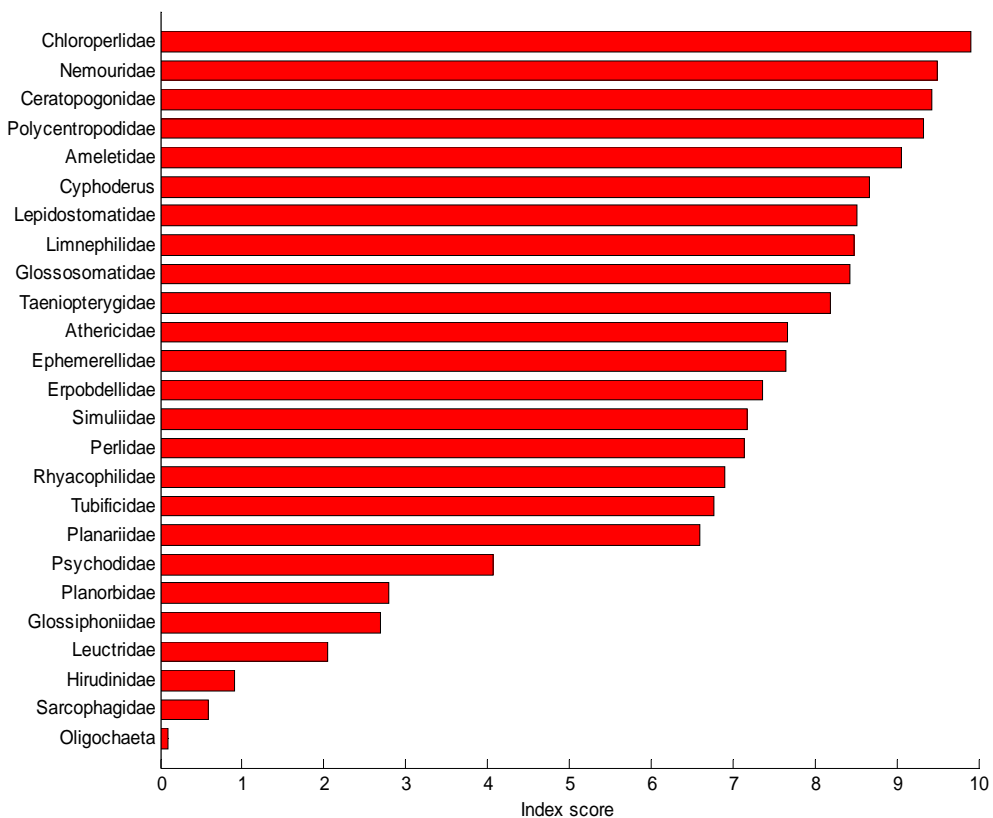


Fig. 6.11. Index scores of the macroinvertebrate populations selected as an indicator in relatively polluted( $w_2$ ) and clean( $w_1$ ) site as a result of the relative importance evaluation.

Table 6.1. Index and the score of selected macroinvertebrate populations.

| <i>Macroinvertebrate Index</i> | <i>Scores</i> | <i>Macroinvertebrate</i> |
|--------------------------------|---------------|--------------------------|
| $I_{10}$                       | 10.0          | Chloroperlidae           |
|                                |               | Nemouridae               |
|                                |               | Ceratopogonidae          |
|                                |               | Polycentropodidae        |
|                                |               | Ameletidae               |
| $I_9$                          | 9.0           | Cyphoderus               |
|                                |               | Lepidostomatidae         |
|                                |               | Limnephilidae            |
|                                |               | Glossosomatidae          |
|                                |               | Taeniopterygidae         |
| $I_8$                          | 8.0           | Athericidae              |
|                                |               | Ephemerellidae           |
|                                |               | Erpobdellidae            |
|                                |               | Simuliidae               |
|                                |               | Perlidae                 |
| $I_7$                          | 7.0           | Rhyacophilidae           |
|                                |               | Tubificidae              |
|                                |               | Planariidae              |
| $I_6$                          | 6.0           |                          |
| $I_5$                          | 5.0           | Psychodidae              |
| $I_4$                          | 4.0           |                          |
| $I_3$                          | 3.0           | Planorbidae              |
|                                |               | Glossiphoniidae          |
| $I_2$                          | 2.0           | Leuctridae               |
| $I_1$                          | 1.0           | Hirudinidae              |
|                                |               | Sarcophagidae            |
|                                |               | Oligochaeta              |

Table 6.2. Comparison of average index scores with average factor scores for the sampling sites.

| <i>Group</i> | <i>Site</i> | <i>Factor score</i> | <i>Estimated factor score</i> | <i>Index score</i> |
|--------------|-------------|---------------------|-------------------------------|--------------------|
| $w_1$        | ReT         | -1.15715            | -1.42943                      | 8.62920            |
|              | DUK         | -1.14136            | -1.35345                      | 8.45996            |
|              | OCU         | -1.07631            | -1.30293                      | 8.34744            |
|              | ReA         | -0.96474            | -1.17178                      | 8.05532            |
|              | DAG         | -0.92740            | -1.12135                      | 7.94300            |
|              | FaC         | -0.89095            | -0.81836                      | 7.26816            |
|              | LiT         | -0.84828            | -0.67235                      | 6.94296            |
|              | ReK         | -0.8171             | -0.98071                      | 7.62975            |
|              | FaA         | -0.74244            | -0.84997                      | 7.33858            |
|              | ONS         | -0.72474            | -0.58268                      | 6.74325            |
|              | DDK         | -0.63792            | -0.73298                      | 7.07799            |
|              | DKS         | -0.62532            | -0.81806                      | 7.26750            |
| $w_2$        | LiK         | -0.59294            | -0.83851                      | 7.31303            |
|              | InJ         | -0.40125            | -0.41128                      | 6.36149            |
|              | HJD         | 0.10559             | -0.52956                      | 6.62493            |
|              | Duk         | 0.49665             | 0.77381                       | 3.72196            |
|              | Dag         | 1.04063             | 0.67425                       | 3.94371            |
|              | Ddk         | 1.26097             | 1.56349                       | 1.96314            |
|              | Dkh         | 1.34811             | 1.57176                       | 1.94471            |
|              | Dks         | 1.35504             | 1.46312                       | 2.18669            |
| SSI          | 1.41726     | -0.69796            | 7.00000                       |                    |



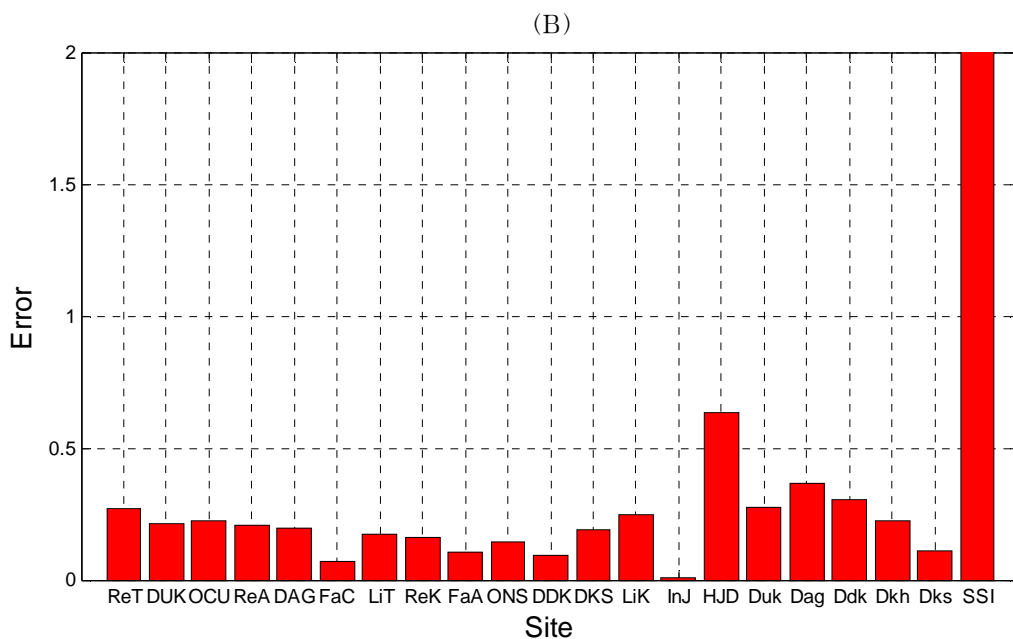
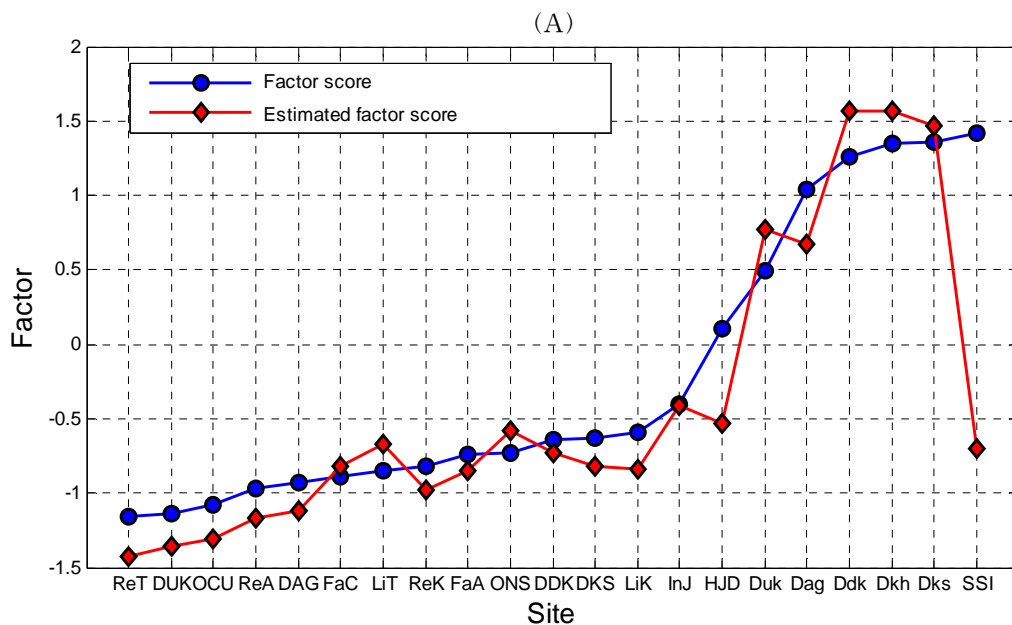


Fig. 6.12. An assessment of water quality using macroinvertebrate indices. (A) Comparison of average scores between factor and estimated factor evaluated by using macroinvertebrate index in each sampling site (B) Estimation error of the factor score in each sampling site.

#### 6.2.4. 유효성 검증

본 연구에서 개발된 생물지수의 유효성을 검증하기 위해 기존에 개발되어 있는 BMWP 점수와 비교하였다. 6.1.3절에서 선정된 대형무척추동물 지수는 BMWP와 동일한 방법으로 산정되었다. BMWP 점수는 각 대형무척추동물에 대해 부여된 내성지수들의 합으로 나타낸다. 이에 본 연구에서 제시한 대형무척추동물 지수를 BMWP의 대형무척추동물의 내성지수에 상응하는 것으로 설정하여 각 지점에서의 지수의 합 계산하여 BMWP 점수와 비교하였다. Fig. 6.13에서 보는 바와 같이 본 연구에서 개발된 지수를 이용한 각 지점의 평가결과는 BMWP를 이용한 평가결과와 유사함을 알 수 있다. 전반적으로 BMWP 점수보다 낮은 양상을 보이거나 오염도를 반영하는 패턴이 유사하였다. 이는 본 연구에서 개발된 생물지수의 이용 가능성을 시사하는 것이다.

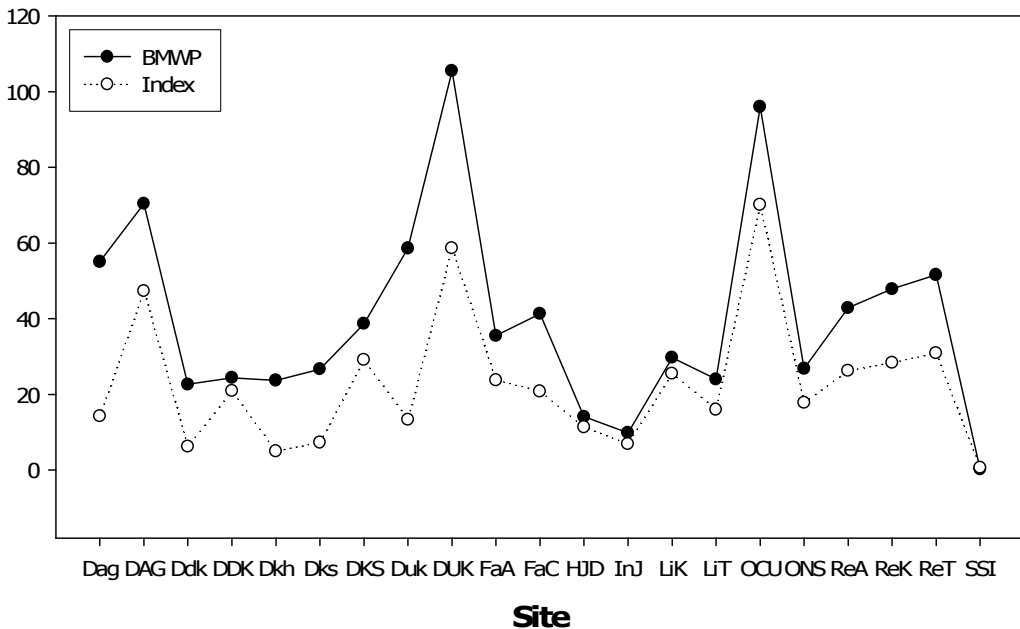


Fig. 6.13. Efficacy assessment of the developed macroinvertebrate index in comparison with the BMWP index.

### 6.3. 결론

본 장에서는 미생물의 지수화 방법을 보다 일반화하여 다른 분류군에 확장·적용함으로써 각 분류군의 수질상태에 대한 특성을 잘 반영하는 통합지수를 개발하기 위하여 제안된 방법을 대형무척추동물에 적용하여 유효성을 검토하였다. 이러한 지수화 방법의 타당성을 검토하기 위해 각 샘플 지역별로 지수를 통해 추정된 요인점수와 요인분석으로부터 추출한 요인의 요인점수를 비교분석하였다. 또한 실제 필드에 적용되고 있는 BMWP지수를 이용한 환경평가 방법을 통해 제안된 방법과 비교분석을 통해 유효성을 평가하였다.

그 결과 대형무척추동물의 지수를 이용한 수질평가는 미생물의 지수를 이용한 평가결과에 비해 전반적으로 우수한 결과를 보였으며 BMWP지수를 이용한 환경평가결과도 전반적으로 유사한 성향을 보이는 것으로 판명되었으며 이를 통해 본 장의 결론을 종합적으로 정리하면 다음과 같다.

1. 확률분석 기법을 이용하여 산정한 대형무척추동물 지수를 이용하여 환경을 평가한 결과 SSI 지점과 같이 심하게 오염된 지점을 제외한 범위에서는 대체적으로 좋은 예측결과를 얻을 수 있었다. 이는 빈영양 또는 극히 오염된 환경에서는 대형무척추동물의 서식이 어려워 해당 환경에 대한 데이터가 없기 때문인 것으로 사료된다.
2. 기존의 BMWP와의 비교를 통해 본 연구에서 개발된 미생물 및 대형무척추동물 지수의 이용가능성을 확인할 수 있었다.
3. 보다 효율적인 수질예측 지수를 얻기 위해서는 오염지역의 분류 그룹을 증가시켜 대형무척추동물의 서식 특성을 파악하는 것이 필요하며, 타 분류군과 상보적으로 적용함으로써 대형무척추동물만으로는 이루어 질수 없는 환경에 대한 평가가 가능할 것으로 사료된다.

## 제 7장 통합지수 개발

1980년대 중반이후로 환경을 관리하고 측정하는 자료로서 생태계 건강도 관리 및 평가에 관한 새로운 수질평가 방법에 관심을 가지게 되었다. 수중 생태계에 영향을 미치는 요인은 화학적 오염뿐만 아니라 다양한 환경변화로 인해 크게 손상되기 때문에 화학적 평가만으로는 어려움이 있다. 이를 보완하기 위해 생물지표를 평가법들이 도입되고 있으며, 나아가 지속적 생태계 관리를 위한 통합수질지표개발을 위한 생태정보 연구가 수행되고 있다.

생물통합지수를 이용한 평가법은 어류를 이용한 IBI가 제안된 이후 대형무척추동물을 이용한 IBI 개발로도 이어져 활발한 연구가 이루어지고 있다. 생물통합지수를 이용한 수질평가는 미국을 시작으로 유럽연합 및 일본을 비롯한 여러 나라에서 적용성을 검토하는 단계이다. 이에 수환경에 관여하는 생물 및 미생물의 제반요인을 내재시키어 종합적으로 나타내어 다면적이며 입체적인 하천 생태계를 표출할 수 있는 통합수질지수의 개발이 요구된다. 그러나 현재까지 대부분의 연구가 어류, 조류, 대형무척추동물 및 미생물 군집별로 독립적으로 연구가 진행되어온 바 생태적 수질평가 및 관리를 위해서는 복잡한 생태환경조건에서 이루어지는 이들 군집의 구조와 기능을 통합하는 생태적 지표 및 지수 개발이 필요하다.

이에 본 연구에서는 생태계의 질을 종합적, 지속적으로 평가할 수 있는 통합수질지수를 개발하고자 하였다. 이를 위해 환경요인과 생물요인을 통합적으로 분석하여 지수를 개발하고자 하였다. 주요 환경요인으로는 수온, pH, 전기전도도, DO, BOD, 탁도, 수심, 유속, COD, TOC, T-P, T-N 및  $\text{NO}_3^-$ -N 등을 분석하였고, 생물요인으로는 조류(생산자), 대형무척추동물(소비자) 및 미생물(분해자)의 군집구조를 분석하였다.

다차원의 환경요인을 요인분석 기법을 통해 하나의 주요인으로 축소하였고 이를 토대로 다분류군의 출현패턴을 확률적 방법을 통해 분석하였다. 이를 통해 해당 환경에 대한 주요한 정보를 제공할 수 있는 생물종들을 효과적으로 선별하고, 선별된 생물종들의 중요도를 지정하기 위해 학습 및 평가 시스템을 구축하여 반복적 모의실험을 통한 각 생물종의 상대적 중요도를 획득하였다. 이로부터 지수를 선정하고 환경평가를 시행하였다.

## 7.1. 실험방법

본 절에서는 통합수질지수를 개발하기 위해 조류, 대형무척추동물 및 미생물 등 세 분류군을 종합적으로 고려하고자 하였다. 이를 위해 세 분류군의 군집분석 데이터를 응용하여 지수를 선정하고 환경 평가를 수행하였다.

환경 데이터 및 다분류군 군집자료는 제 5장에서 제안한 바와 동일한 절차에 따라 분석을 수행하였다. 각 환경변수를 대표할 수 있는 공통성분을 추출하고, 다분류군과 공통요인간의 관계를 규명하기 위해 공통요인에 대한 데이터를 두 개의 그룹으로 분류하였다. 각각의 그룹에 대한 다분류군 출현밀도를 확률모델로 근사한 후 이를 비교분석하여 다분류군 내의 각 생물종을 특성화함으로써 그 출현패턴을 최우추정법을 통해 추정된 가우시안 파라미터를 이용하여 클래스-조건부 확률밀도함수로 근사하여 분포특서에 따라 4그룹으로 분류하였다. 그리고 PNN을 이용한 상대적 중요도 평가시스템을 통해 분류군에 관계없이 수질 예측력이 상대적으로 우수한 종들을 선정하였고, 선정된 생물종들 각각에 대해 상대적 중요도를 부여하였다. 주어진 상대적 중요도에 따라 정렬된 변수들이 수질평가가 가능하도록 요인공간으로 사상시킨 후 선형 스케일링 함수를 통해 1에서 10의 범위를 가지는 지수를 부여하였다.

## 7.2. 결과 및 고찰

### 7.2.1. 미생물-대형무척추동물 통합지수

본 절에서는 미생물과 대형무척추동물 군집구조를 이용하여 통합지수를 개발하였다. 중요도 평가를 통해 부여된 지수는 Table 7.1에 나타내었다. 환경평가를 위한 지표로 선정된 미생물은 총 30종이며, 대형무척추동물은 25종이 선정되었다. 대부분의 대형무척추동물은 높은 지수점수가 부여되었으며, 미생물은 낮은 지수점수가 부여되었다. 가장 높은 지수점수를 가지는 생물종은 Chloroperlidae, Ceratopogonidae, Cyphoderus, Nymphidae, Polycentropodidae로 모두 대형무척추동물이 선정되었고 상대적 청정지역에서 중요도가 가장 높은 생물이 된다. 가장 낮은 점수를 가지는 생물종은 미생물의 경우 *Methylobacterium* sp., *Arthrobacter* sp., *Hymenobacter* sp., *Klebsiella* sp., *Herbaspirillum* sp., *Duganella* sp. 및 *Bdellovibrio* sp., 등이, 대형무척추동물의 경우 Sarcophagidae와 Oligochaeta이 선정되어 상대적 오염지역에서 중요도가 가장 높은 생물이 된다. 또한 지수점수 5 이하로는 선정된 대형무척추동물이 거의 없고 대부분이 미생물들이 선정되었다. 전반적으로 높은 지수점수에는 대형무척추동물이, 낮은 지수점수에는 미생물이 선정되었다. 이는 오염도를 평가함에 있어 청정지역에서는 대형무척추동물, 오염지역에서는 미생물이 중요한 역할을 하는 것으로 판단할 수 있다.

각 생물종에 대해 부여된 지수를 이용하여 환산된 요인점수와 각 지점의 실제 요인점수의 비교를 통해 환경평가 유효성을 검토하였다(Fig. 7.1). 그 결과 각 분류군에 대해 개별적으로 산정된 지수를 이용한 예측결과에 비해 우수한 예측결과가 얻어졌다. 이는 두 분류군의 통합을 통해 단일 분류군이 반영하지 못하던 환경에 대해 상호보완 해줌으로서 예측오차를 크게 향상시킬 수 있었던 것으로 사료된다. 미생물의 경우 중간오염지역에 대한 평가 지표가 부족하였고, 대형무척추동물의 경우 오염이 심하거나 빈영양 지역에서의 평가가 어려움이 있었으나 두 분류군을 통합함으로써 예측결과를 높일 수 있었던 것으로 사료된다.

Table 7.1. Scores for the integrated indices based upon the multi-taxa (microorganisms and macroinvertebrate).

| <i>Integrated Index</i> | <i>Scores</i> | <i>Organisms</i>   |  |
|-------------------------|---------------|--|--|
|                         |               | <i>Microorganisms</i>  | <i>Macroinvertebrate</i>   |
| $I_{10}$                | 10.0          |  | Chloroperlidae<br>Ceratopogonidae<br>Cyphoderus<br>Nemouridae<br>Polycentropodidae   |
| $I_9$                   | 9.0           | <i>Stenotrophomonas maltophilia</i><br>Uncultured Verrucomicrobiales<br><i>Lactobacillus delbrueckii</i>   | Ameletidae<br>Glossosomatidae<br>Limnephilidae<br>Athericidae  |
| $I_8$                   | 8.0           | <i>Pandoraea</i> sp.<br>Uncultured <i>Nitrospira</i> sp.<br><i>Ralstonia</i> sp.<br><i>Propionibacterium</i> sp.<br><i>Rhodobacter</i> sp.   | Taeniopterygidae<br>Lepidostomatidae<br>Ephemerellidae<br>Simuliidae<br>Erpobdellidae<br>Rhyacophilidae<br>Perlidae<br>Planariidae |
| $I_7$                   | 7.0           | Uncultured Bacteroidetes<br><i>Lactosphaera</i> sp.<br><i>Acidobacteria bacterium</i>  | Tubificidae<br>Tipulidae<br>Heptageniidae  |
| $I_6$                   | 6.0           |  |  |
| $I_5$                   | 5.0           | <i>Chryseobacterium</i><br><i>Planorbidae</i> sp.  |  |
| $I_4$                   | 4.0           | <i>Bacillus</i> sp.<br><i>Sphingobium</i> sp.  |  |
| $I_3$                   | 3.0           | <i>Janthinobacterium</i> sp.<br><i>Exiguobacterium</i> sp.   | Glossiphoniidae<br>Leuctridae  |
| $I_2$                   | 2.0           | <i>Dechloromonas</i> sp.<br><i>Comamonas terrigena</i><br><i>Lysobacter</i> sp.<br><i>Caulobacter</i> sp.<br><i>Brevundimonas</i> sp.<br><i>Rhodoplanes</i> sp.                            | Hirudinidae  |
| $I_1$                   | 1.0           | <i>Methylobacterium</i> sp.<br><i>Arthrobacter</i> sp.<br><i>Hymenobacter</i> sp.<br><i>Klebsiella</i> sp.<br><i>Herbaspirillum</i> sp.<br><i>Duganella</i> sp.<br><i>Bdellovibrio</i> sp. | Sarcophagidae<br>Oligochaeta   |

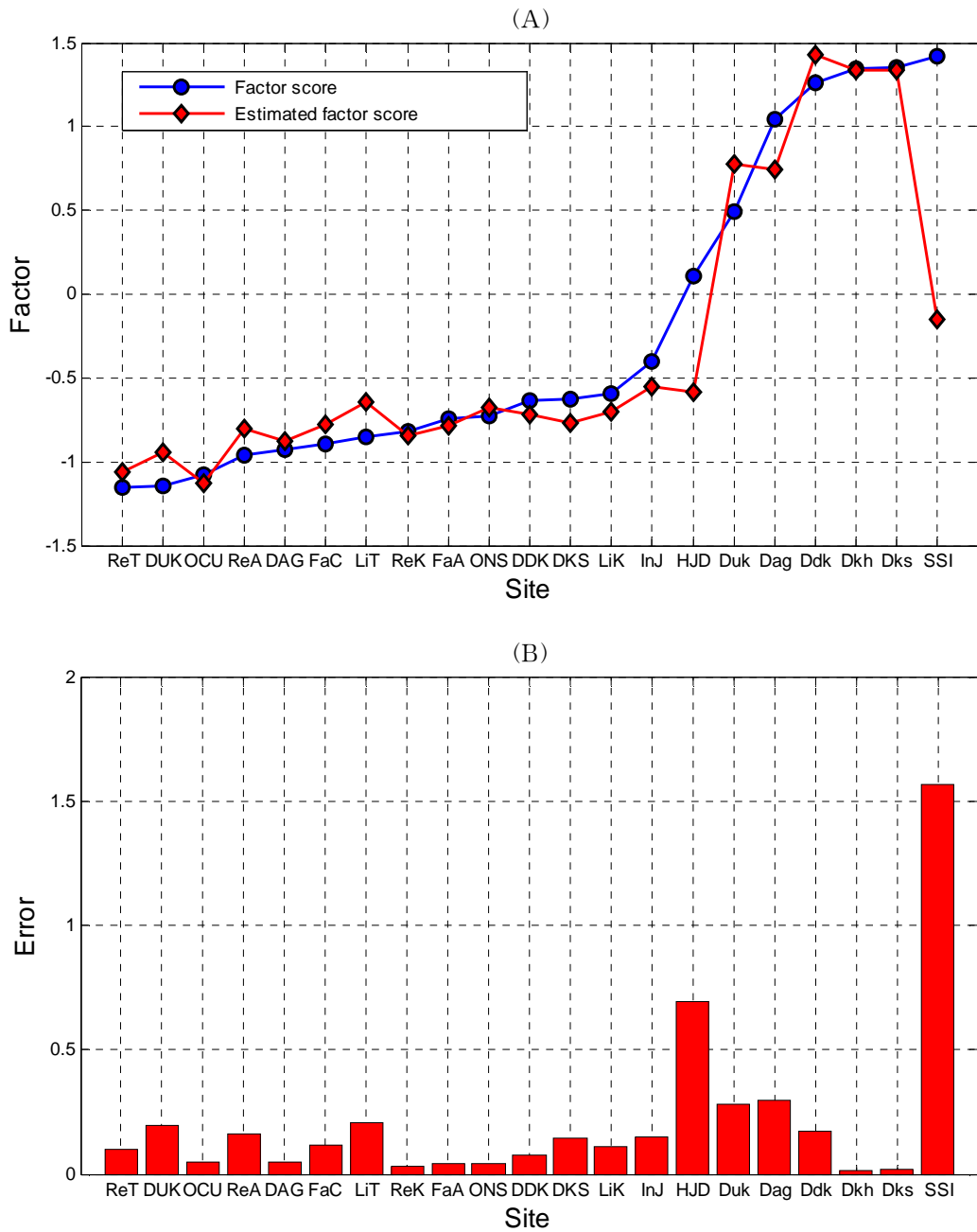


Fig. 7.1. An assessment of water quality using the integrative water quality index based upon the multi-taxa (microorganisms and macroinvertebrates). (A) Comparison of average scores between factor and estimated factor evaluated by using integrative index in each sampling site; (B) Estimation error of the factor score in each sampling site.



## 7.2.2. 미생물-대형무척추동물-조류 통합지수

미생물, 대형무척추동물, 조류 군집구조를 이용하여 통합지수를 개발하였다. 각 생물종의 상대적 특성을 종합적으로 고려하여 상대적인 중요도를 평가하였고, 이를 통해 수질의 오염도를 평가할 수 있는 생물종을 선별하였다. 선별된 생물종들에 평가된 상대적 중요도에 따라 지수를 부여하였다. 중요도 평가를 통해 부여된 지수는 Table 7.2와 같다. 환경평가를 위한 지표로 선정된 생물종은 미생물 33종, 대형무척추동물 24종, 조류 14종이 선정되었다. 전반적으로 대형무척추동물과 조류는 높은 지수점수가 부여되었으며, 미생물은 전반적으로 분포하나 낮은 지수점수에 많이 선정되었다. 특히 조류의 경우는 높은 지수점수에 편중되어 지수점수 7이하로 선정된 종은 없었다.

가장 높은 지수점수를 가진 생물종은 대형무척추동물의 Chironomidae, Culicidae 및 Naididae와 조류의 Gomphonema가 선정되었다.

각 생물종에 대해 부여된 지수를 이용하여 환산된 요인점수와 각 지점의 실측 요인점수의 비교를 통해 환경평가 유효성을 검토하였다(Fig. 7.2). 제시된 바와 같이 비교적 좋은 예측결과를 얻었으나 두 분류군(미생물과 대형무척추동물)을 통해 제시된 지수의 평가결과보다 오차가 크게 나타났다. 그러나 비교적 오염도가 낮은 ReT, OCU, DAG 및 ReK 지점의 경우에는 세 분류군을 이용한 지수가 더 좋은 예측결과를 얻었다. 이는 타 분류군에 비해 조류 데이터가 많은 missing data 가지고 있었던 데 기인한 것으로 사료된다. 본 연구에서 수집된 다분류군 군집데이터 중 미생물과 대형무척추동물 군집조사 데이터는 전 조사시기 및 전 지점에 대한 데이터가 모두 수집된 반면 조류 데이터의 경우 2002년 1월부터 2004년 1월까지 Duk, Ddk, Dkh, Dks 및 Dag 지점에 대한 19개월의 데이터가 없어 상대적 오염지역인 위 5지점은 물론 오염도가 높은 지점들에 대한 예측결과가 좋지 않은 것으로 사료된다. 이는 보완하기 위해서는 다양한 오염도에서의 조류 군집 데이터를 보충함으로써 향상시킬 수 있을 것으로 사료된다.

Table 7.2. Score for the integrated index based upon multi-taxa (microorganism, macroinvertebrate and algae).

| Integrated Index | Scores | Organisms  |   |  |
|------------------|--------|--|---|--|
|                  |        | Microorganisms   | Macroinvertebrate   | Algae  |
| $I_{10}$         | 10.0   |  | Chironomidae<br>Culicidae<br>Naididae   | Gomphonema   |
| $I_9$            | 9.0    | <i>Lactobacillus delbrueckii</i><br>Uncultured <i>Verrucomicrobiales</i><br><i>Stenotrophomonas maltophilia</i><br>Uncultured <i>Nitrospira</i>  | Capniidae<br>Glossiphoniidae<br>Pleuroceridae<br>Libellulidae   | Chroococcus  |
| $I_8$            | 8.0    | <i>Pandoraea</i> sp.<br><i>Propionibacterium</i> sp.<br><i>Ralstonia</i> sp.<br>Uncultured Bacteroidetes<br><i>Rhodobacter</i> sp.   | Hydroptilidae<br>Astacidae<br>Pyralidae<br>Oligochaeta<br>Sarcophagidae<br>Enchytraeidae<br>Ephydriidae | Veliidae<br>Spirulina<br>Euglena<br>Crytomonas<br>Anabaena<br>Surirella<br>Pediastrum<br>Klebsormidium<br>Microcystis<br>Geminella |
| $I_7$            | 7.0    | <i>Lactosphaera</i> sp.<br><i>Acidobacteria bacterium</i><br><i>Flavobacterium</i> sp.<br><i>Erythrobacter</i> sp.   | Stratiomyiidae<br>Trichoptera<br>Physidae<br>Helodidae  | Navicula<br>Cyclidiopsis   |
| $I_6$            | 6.0    |  |   |  |
| $I_5$            | 5.0    |  |   |  |
| $I_4$            | 4.0    | <i>Chryseobacterium</i> sp.<br><i>Sphingobium</i> sp.<br><i>Bacillus</i> sp.<br>alpha proteobacterium  | Planariidae   |  |
| $I_3$            | 3.0    | <i>Janthinobacterium</i> sp.   | Gastropoda  |  |
| $I_2$            | 2.0    | <i>Dechloromonas</i> sp.<br><i>Exiguobacterium</i> sp.<br><i>Lysobacter</i> sp.<br><i>Rhodoplanes</i> sp.<br><i>Comamonas terrigena</i>  | Leptophlebiidae<br>Hirudinea  |  |
| $I_1$            | 1.0    | <i>Brevundimonas</i> sp.<br>beta proteobacterium<br><i>Caulobacter</i> sp.<br><i>Methylobacterium</i> sp.<br><i>Klebsiella</i> sp.<br><i>Herbaspirillum</i> sp.<br><i>Arthrobacter</i> sp.<br><i>Hymenobacter</i> sp.<br><i>Bdellovibrio</i> sp.<br><i>Duganella</i> sp. | Rhyacophilidae<br>Odontoceridae   |  |

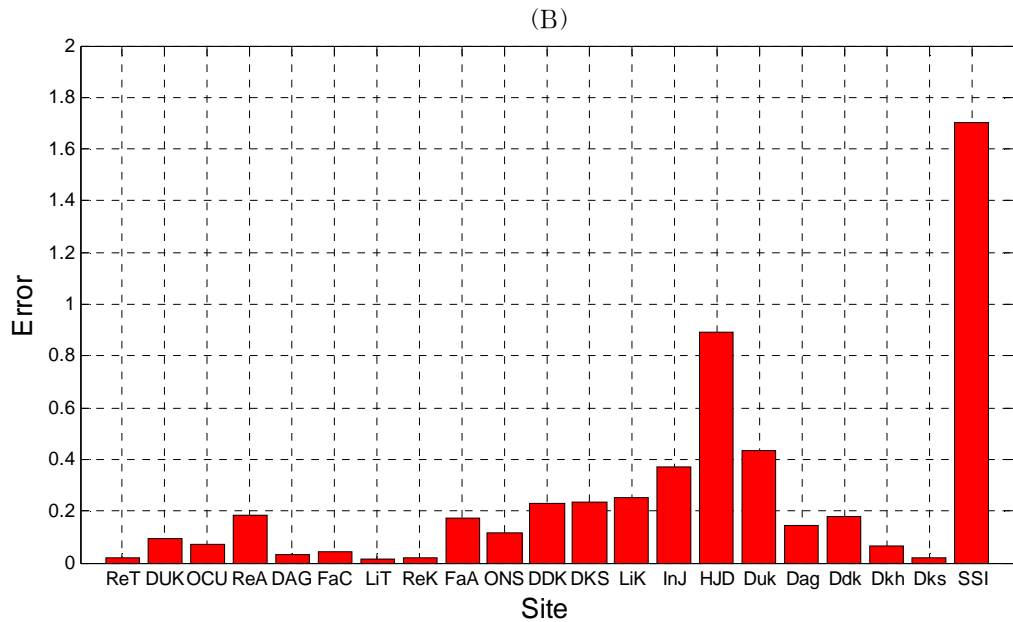
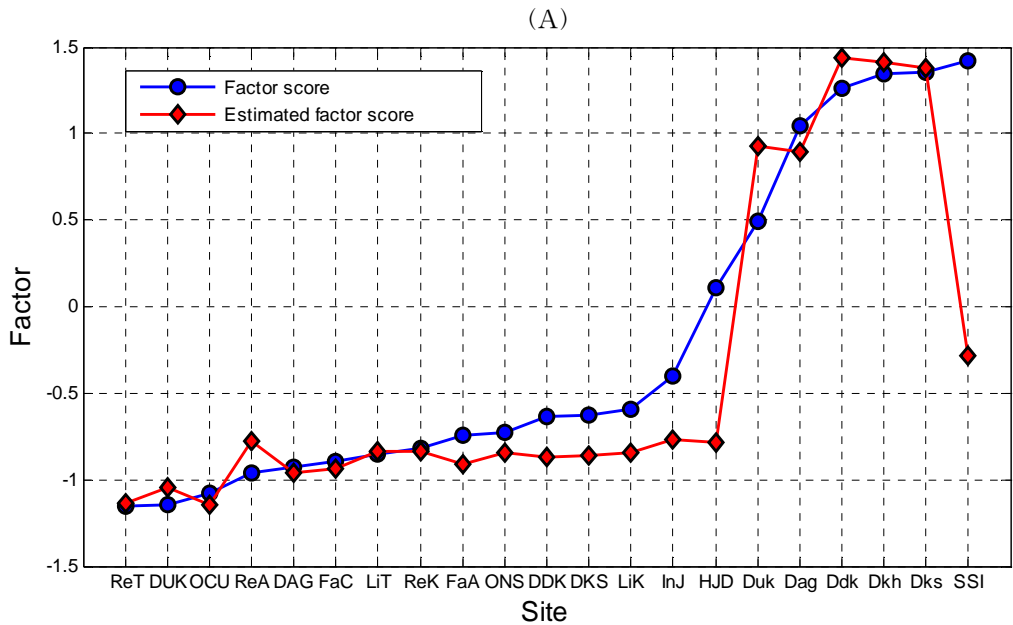


Fig. 7.2. An assessment of water quality using the integrative water quality index based upon the multi-taxa (microorganisms, macroinvertebrates and algae). (A) Comparison of average scores between factor and estimated factor evaluated by using integrative index in each sampling site; (B) Estimation error of the factor score in each sampling site.

### 7.2.3. 통합지수 평가

본 연구에서는 효과적인 수질평가를 위한 방법으로 담수생태계의 수질에 결정적인 원인을 제공하는 조류, 무척추저서생물 및 미생물에 대한 개별적 인과관계에 그치지 않고 생산자-소비자-분해자의 관점에서 수질상태를 잘 반영하는 개체군들의 지수화를 목적으로 하였다. 이를 위해서는 각 분류군의 군집특성을 우선적으로 파악하는 것이 중요하였으며 이를 기반으로 다음의 4가지 경우에 대한 수질평가방법을 모두 고려하여 효과적인 통합지수를 얻고자 하였다.

- (1) 미생물 지수를 이용한 수질평가 (MI)
- (2) 대형무척추동물 지수를 이용한 수질평가 (MA)
- (3) 미생물+대형무척추동물 지수를 이용한 수질평가 (MI+MA)
- (4) 미생물+대형무척추동물+조류 지수를 이용한 수질평가 (MI+MA+AL)

제 5장에서는 (1)항목의 미생물에 대한 지수 및 환경평가를 우선적으로 다루었고 (2)의 항목은 제 6장에서는 다름으로써 미생물의 지수화 방법을 대형 무척추저서생물에 적용하여 본 논문에서 제안한 지수화 방법의 효용성을 입증하였으며 다른 개체군에의 적용가능성을 보여 주었다. 본 장에서는 (3)과 (4)항목을 동시에 다루었으며 미생물과 대형무척추동물을 통합한 지수를 이용하여 보다 정확한 환경평가가 가능함을 확인하였다. 그러나 조류를 포함한 통합지수는 오히려 환경평가의 정확도가 떨어졌는데 이는 상대적 오염지역에서의 조류데이터가 전무하였기 때문이었다. 이는 본 연구에 이용된 생물자료 중 조류 군집구조 데이터가 타 분류군에 비해 부족했기 때문인 것으로 사료된다.

4 종류의 통합지수를 활용하여 수질평가의 결과를 최종 검토하기 위하여 각각의 경우에 대하여 각 샘플 지역의 절대오차를 평균하여 이들을 비교하였다 (Table 7.3). 실험결과에서 4가지 생물지수 모두에서 SSI 지점의 오차는 상대적으로 매우 크기 때문에 오차비교표에서는 이 지역을 제외하였다. 평균오차가 가장 큰 통합지수는 미생물 지수만을 이용한 경우로 판명되었다. 또한 MA의 경우 MI에 비해 오차는 적었으나 두 분류군의 통합지수에 비해 오차가 컸다. 가장 좋은 성능을 보이는 통합지수는 MI+MA의 경우로 나타났다. 두 분류군을 이용한 통합지수 MI+MA가 세 분류군을 이용한 통합지수 MI+MA+AL보다 좋

은 나은 예측 성능을 보인 것은 조류의 데이터가 상대적 오염지역에서 대부분 결측값을 가지고 있었기 때문에 오염지역으로 갈수록 수질상태를 잘 반영하지 못했기 때문으로 사료된다. 그러나 청정지역으로 갈수록 MI+MA+AL의 경우 예측오차가 4가지 경우 중 가장 작은 것으로 분석되었다.

Fig. 7.3은 4가지의 통합지수를 이용한 요인예측결과를 각 지점별로 나타낸 것이다. M1의 경우 대부분의 지역에서 가장 현저한 오차를 보였다. 특히 중간 지역의 수질에 대하여 M1만으로는 중간지역을 설명하기 어려운 것으로 보여졌다. MA의 경우 DAG이하의 청정지역과 Dag 이상의 오염지역에서는 다른 경우에 비해 상대적으로 많은 오차를 보이는데 이는 대형무척추동물의 분포특성에 의한 것으로 오염이 심하거나 빈영양 상태의 청정지역에서는 소비자인 대형무척추동물의 종풍부도 및 다양성이 낮은 것에 기인한 것으로 사료된다. 가장 우수한 성능을 보이는 MA+MI의 경우 특히 오염지역에 대한 오차가 다른 경우에 비해 적은 것으로 분석되었다. 이는 상대적으로 오염지역에서는 미생물의 지수가 청정지역에서는 대형무척추저서생물의 지수가 환경평가에 잘 반영되어 상호 보완적인 작용을 하여 M1 및 MA의 경우에 비해 좋은 성능을 보이는 것으로 사료된다. MI+MA+AL의 경우 청정지역에 대해서는 오히려 MI+MA에 비해 더욱 좋은 성능을 보이는 것으로 확인되었다. 이러한 분석결과로부터 오염지역에서의 조류데이터가 충분히 확보 가능하다면 제안된 통합 지수를 이용하여 보다 정확한 환경평가가 이루어 질 것이라 기대된다.

Table 7.3. Comparison of estimation performance for each biological index in terms of IAE (Integral Absolute Error).

|     | Biological indices |          |         |          |
|-----|--------------------|----------|---------|----------|
|     | MI                 | MA       | MI+MA   | MI+MA+AL |
| IAE | 0.28328            | 0.240300 | 0.14707 | 0.16698  |

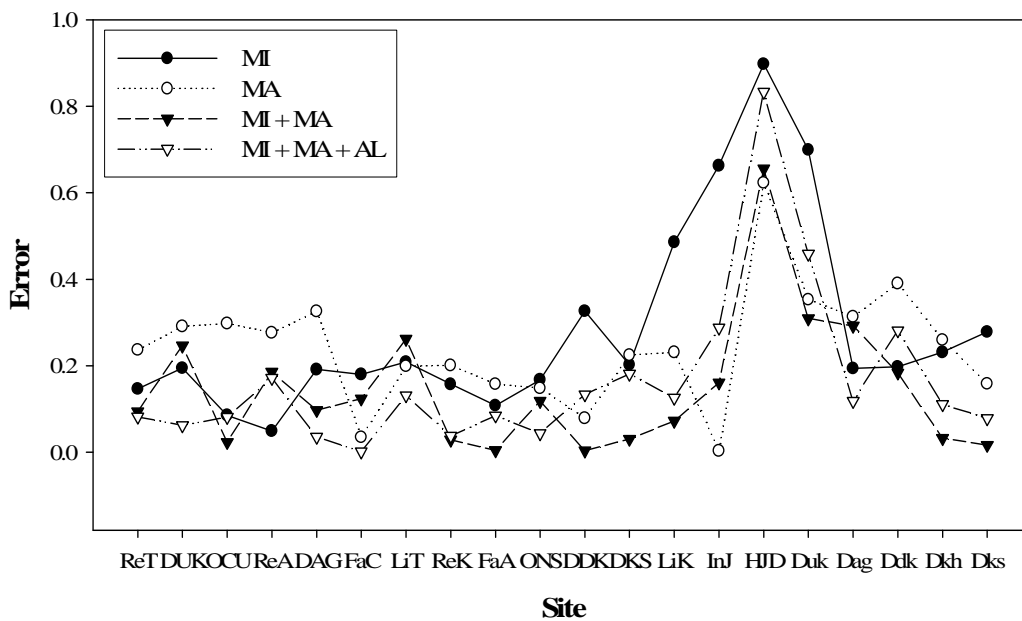


Fig. 7.3. Comparison of the estimation performance for water quality assessment in four cases of the integrative water quality indices. MI: microbial index; MA: macroinvertebrate index; AL: algal index.

## 7.4. 결론

본 연구에서는 효과적인 수질평가를 위한 방법으로 조류, 무척추저서생물 및 미생물을 이용하여 통합수질지수를 개발하였다. 개발된 통합수질지수들의 비교를 통해 다분류군을 이용하여 개발된 지수가 단일 분류군을 이용한 수질평가보다 효과적이며 환경을 잘 반영할 수 있음을 확인하였다. 또한, 각 분류군의 생태 및 서식 특성에 의해 오염도 및 생태계의 상황에 따라 군집구조에 차이가 있어 특정 환경의 상황을 각 분류군이 반영할 수 없는 환경에 대해 상호 보완적으로 작용하여 보다 정확한 환경평가가 이루어 질 수 있을 것이라 기대된다. 이를 통해 본 장의 결론을 종합적으로 평가하면 다음과 같다.

1. 다분류군을 이용하여 통합지수를 선정하여 환경평가를 수행한 결과 단일 분류군만으로 평가한 것에 비해 좋은 예측결과를 얻을 수 있었다. 이는 통합지수가 단일 분류군에 의한 환경평가 능력을 제고하는데 상당히 기여할 수 있음을 시사한다.
2. 오염하천 생물 콤플렉스의 구조와 기능을 생태학적으로 평가, 검토할 수 있는 통합수질지수를 개발하기 위해서는 다양한 환경에서의 다분류군 군집특성 연구가 필요하며 각 분류군을 보다 정량적으로 측정할 수 있는 방법이 개발되어야 할 것으로 사료된다.

## 제 8장 종합결론

본 연구에서는 효과적인 수질평가를 위한 방법으로 조류, 무척추저서생물 및 미생물에 대한 개별적 인과관계에 그치지 않고 생산자-소비자-분해자의 관점에서 수질상태를 잘 반영하는 생물들을 이용하여 통합수질지수를 개발하였다. 각 분류군의 군집특성을 우선적 파악하여 이를 기반으로 효과적인 통합지수를 얻을 수 있었다. 개발된 통합수질지수들의 비교를 통해 다분류군을 이용하여 개발된 지수가 단일 분류군을 이용한 수질평가보다 효과적이며 환경을 잘 반영할 수 있음을 확인하였다.

1. 생물군집(미생물, 조류 및 저서생물) 분석결과 이화학적 환경자료와 각 생물군의 군집들은 대체로 오염 지역의 특성에 따른 SOM 분류가 이루어짐을 관찰하였다. 특히 저서 생물의 경우 청정지역과 오염지역 종류들 간에 뚜렷이 구분되는 양상을 보였다. 즉 오염 지역에 따른 생물군집의 양상을 구분할 수 있게 된 것이다. 미생물 및 조류의 군집의 경우는 지역보다 계절적 영향에 더 민감한 것으로 나타나서 분류군간의 특성을 인정해야 할 필요가 있는 것으로 보인다.
2. 미생물 지수를 개발하기 위해 PCR-DGGE 기법을 이용하여 미생물의 상대적인 밀도를 측정하여 미생물 지수 선정에 이용하였다. 개발된 미생물 지수는 수생태계의 오염도를 대체적으로 잘 반영하여 수생태계의 오염도 평가에 활용할 수 있음을 확인하였다. 그러나 SSI, InJ, Duk 및 일부 지점에서 오차범위가 크게 평가되었다. 이는 선정된 일부 미생물로는 이 지점의 환경을 설명할 수 없었기 때문으로 사료된다. 이는 좀 더 다양한 미생물 종의 분석을 통해 생물지수를 확대함으로써 향상시킬 수 있을 것으로 사료된다.
3. 미생물 밀도조사를 위해 사용된 PCR-DGGE 조건(16S rRNA 유래 primers, PCR 조건, DGGE gel에서의 PCR product separation 조건 등)은 매우 양호한 미생물 군집 분석 data의 획득에 기여한 것으로 사료된다.



다. 그러나 보다 신뢰성 있는 데이터 확보를 위해 normalization 및 internal standard DNA를 첨가 기법의 도입이 요구된다.

4. 본 연구에서 개발된 지수를 이용하여 환경을 평가한 결과 대체적으로 좋은 예측결과를 얻을 수 있었다. 각 분류군에 대해 개별적으로 적용하여 지수를 산정하고 환경을 평가한 결과 각 분류군 마다 특정 오염지역에 대해 오차가 크게 발생하는 것을 발견하였다. 미생물지수의 경우는 중간오염수준 지역에 대한 오차가 높게 나타나고 대형무척추동물의 경우는 오염이 심한 지역에서의 오차가 크게 나타났다. 이는 두 분류군을 통합하여 지수를 산정함으로써 오차를 줄일 수 있었다.
5. 확률분석을 이용하여 개발된 대형무척추동물 지수와 기존 개발 지수와의 비교를 통해 개발된 지수의 유효성을 검토하였다. 개발된 대형무척추동물 지수는 BMWP를 이용한 환경평가 결과와 상당히 유사한 결과를 얻을 수 있었다. 이는 본 연구에서 개발된 생물지수의 환경평가에의 적용 가능성을 시사한다.
6. 각 분류군을 대표할 수 있는 지수를 선정하여 이를 통해 개발되는 통합수질지수는 생산자(부착조류), 소비자(저서생물), 분해자(미생물)의 관점에서 생태계 질을 평가할 수 있을 것으로 사료된다.
7. 본 지수는 평가에 주안점을 둔 환경 항목들이 일반적인 수질평가 항목임을 감안한다면 특정 오염물(중금속, 독성 및 난분해성 유기화합물 등)에 의한 영향을 표출하기에는 어려움이 있을 것으로 추정된다. 특정오염물에 의한 환경영향을 평가하기 위해서는 보다 다양한 오염환경에 대한 시료를 늘려 조사하고, 특정오염물 관련 생물군에 대한 조사 검토가 필요할 것으로 사료된다.

## 제 9장 참고문헌

- 곽인실, 김좌관, 전태수 (2003) 인공신경망을 이용한 도시하천의 저서성 대형무척추동물군집 유형성 연구. Korean J. Limnol. 36(1): 29-37.
- 김대일, 이종호, 박희경, 이동률 (2002) 유역 통합 수질 관리를 위한 지속가능개발 지표의 개발. 상수도학회지. 16: 322-332.
- 김미아, 김재관, 조경덕 (2007) 다변량분석법을 이용한 금강 유역의 수질오염특성 연구. J. Korean Soc. Water Quality. 23: 161-168.
- 김상종 (1997) G-7 연구성과세미나 발표집. 국립환경연구원 환경기술개발관리센터.
- 김영주 (2003) 주성분분석에 의한 도시호수의 수질평가에 관한 연구. 환경관리학회지, 9: 197-203.
- 김자현, 안광국 (2004) 대전천에서의 생태학적 건강성 평가. 대한상수도학회 · 한국물환경학회 추계학술발표회 논문집 p. 421
- 김종구 (2002) 통계분석 기법을 이용한 금강수계의 수질평가. 한국환경과학회지. 11: 1281-1289.
- 나철호, 최충길, 백순기 (1986) 한국산 수서곤충의 환경오염물질에 대한 내성 및 감수성연구. 한육수학회지 19: 109-125.
- 박종화 (1997) 국내 여건에 맞는 자연형하천공법의 개발. 한국건설기술연구원, 환경부.
- 배연재, 박선영, 윤일병, 박재홍, 배경석 (1996) 왕숙천 준설구간의 저서성 대형무척추동물 군집변동. 한국육수학회지. 29(4): 251-261.
- 서영희, 박상옥 (1982) 경산 남천 수계의 생물학적 수질판정. 한국생태지. 5: 46-53.
- 신성교, 박청길, 송교욱 (1998) 주성분 분석을 이용한 낙동강 하구 해역의 수질평가. 한국환경관리학회지. 7: 171-176.
- 염동혁, 안광국, 홍영표, 이성규 (2000) 어류군집을 이용한 금호강의 생물보전지수(Index of Biological Integrity, IBI) 평가. Korean J. Environ. Biol. 18(2): 215-226.

- 이영준 (2002) 요인분석의 이해. 석정.
- 이재연, 장하나, 안광국 (2004) 유등천에서의 수질 양상 및 생물학적 건강도 평가. 대한상수도학회 · 한국물환경학회 추계학술발표회 논문집. p. 405.
- 이현엽, 문경일 (1999) MATLAB을 이용한 퍼지-뉴로. 아진.
- 윤일병, 배연재, 노태호, 이성진, 박재홍 (1998) 수서곤충 군집분석에 있어서 최적표본크기의 결정-방태천 모형. 한국생태학회지. 21(5-1): 409-418.
- 윤일병 (1995) 고려대학교 생물학과 생태학 연구실. 수서곤충검색도설 p. 261.
- 장현정, 이용욱 (2003) 상수원 수질관리를 위한 분변오염 지표세균에 관한 연구. 한국환경위생학회지. 29: 19-27.
- 최성현 (1997) 수질모니터링을 위한 생물경보장치의 시범운영현황, 수질모니터링의 현황과 전망. 국립환경연구원.
- 최지용 (1996) 종합수질지표의 개발. 한국환경기술개발원.
- 최지용, 신은성 (1997) 하천수질관리를 위한 종합수질지표의 개발과 적용. 한국수질보전학회지. 13, 415-425.
- 한국화학연구소. (1995) 환경오염물질의 위험성 확인 및 독성평가에 관한 연구(Ⅱ)-환경위해성 평가 및 관리기술(제2차년도 보고서).
- 환경부 (1978) 수질환경기준(환경보전법 시행규칙).
- 환경부 (1989) 수질오염실태조사지침.
- 환경부 (2006) 수질 및 수생태계 환경기준 5차 개정(환경정책기본법 시행령 개정).
- APHA, AWWA, WPCF (1985) *Standard methods for the examination of water and waste* (16th ed.). Washington D.C.
- Bernard, P., L., Antoine and L., Bernard (2004) Principal component analysis: an appropriate tool for water quality evaluation and management-application to a Tropical Lake system. *Ecological Modeling*. 178: 295-311.
- Bishop, C. M. (2006) *Pattern Recognition and Machine Learning*. Springer.
- Bozzetty, M. and U. H. Schulz (2004) An index of biotic integrity based on fish assemblages for subtropical streams in southern Brazil. *Hydrobiologia*. 529: 133-144.
- Brighnam, A. R., W. U. Brighnam, and A. Gnika (1982) Aquatic insects and

- Oligochaeta of North and South Carolina. Midwest Aquatic Enterprise.
- Brinkhurst, R. O. (1986) Guide to the Freshwater Aquatic Microdrile Oligochaetes of North America. Canadian Special Publication of Fisheries and Aquatic Sciences. 84: 259.
- Brown, R. M. (1970) A Water Quality Index- Do We Ware? Water and Sewage Works. 33-343, 1970.
- Chon, T.-S., Y-S. Park, K-H. Moon, E-Y. Cha (1996) Patterning communities by using an artificial neural network. *Ecol. Model.*, 90: 69-78.
- Chon, T.-S., Y-S. Park and J-H. Park (2000) Determining temporal pattern of community dynamics by using unsupervised learning algorithms. *Ecol. Model.* 132: 151-166.
- Chon, T.-S., Y-S. Park and E-Y. Cha (2000) Patterning of community changes in benthic macroinvertebrates collected from urbanized streams for the short time prediction by temporal artificial neural networks. In: Artificial Neural Networks in Ecology and Evolution (Lek S. and J. F. Guegan eds). Springer-Verlag, Berlin.
- Chon, Tae-Soo, Inn-Sil Kwak, Young-Seuk Park, Tae-Hyung Kim, YooShin Kim (2001) Patterning and short-term predictions of benthic macroinvertebrate community dynamics by using a recurrent artificial neural network. *Ecological Modelling.* 146: 181-193.
- Costanza, R., B. G. Norton and B. D(Eds). Haskell (1992) Ecosystem Health: New Goals for Environmental Management. Island Press, Washinfon, DC, p. 239-256.
- Cummins, K. W. and G. H.Lauff (1969) The influence of substrate particle size on the microditribution of stream macrobenthos. *Hydrobiologia.* 34: 145-181.
- Daily, G. C. (1997) Nature's Services: Societal Dependence on Natural Ecosystem. Island Press, Washington, DC.
- Deiniger, R. A. and Jurate, M. (1971) A Water Quality Index for Public Water Supplies, Unpublished Report. Department of Environmental

- Health. University of Michigan, Ann Arbor, MI, July, 17-35.
- Duda, R. O., P. E. Hart and D. G. Stork (2001) Pattern Classification. 2nd edition. John-Wiley.
- Gu, M.-B., R. J. Mitchell and J.-H. Kim (2000) Continuous monitoring of protein damaging toxicity using a recombinant bioluminescent *Escherichia coli*. ACS Symposium Series. Recent Advances in Chemical sensors and Biosensors for Environmental monitoring. 762: 185-196.
- Gu, M.-B., G. C. Gil and J. H. Kim (1999) A two-stage minibioreactor system for continuous toxicity monitoring. Biosensors & Bioelectronics. 14: 355-361.
- Hastie, T., R., Tibshirani, and J. Friedman (2001) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
- Hecht-Nielsen, R. (1990) Neurocomputing. Addison-Wisley. Sydney. p. 433.
- Hellawell, J.M. (1986) Biological Indicators of Freshwater Pollution and Environmental Management. Elsevier, London. 546pp.
- Hynes, H.B.N. and M.J. Coleman (1968) A simple method of assessing the annual production of stream benthos. *Limnol. Oceanogr.* **13**: 569-573.
- Jorgensen, S. E., S. N. Nielson and H. F. Mejer (1995) Emergy, environ exergy and ecological modeling. Ecological Modeling. 77: 99-109.
- Karr, J. R. (1981) Assessment of biotic integrity using fish communities. Fisheries. 6: 21-27.
- Karr, J. R. (1991) Biological integrity: A long-neglected aspect of water resource management. Ecol. Appl. 1: 66-84..
- Kelly, M. G. and B. A. Whitton (1995) The Trophic Diatom Index: a new index for monitoring eutrophication in rivers. Journal of Applied Phycology. 7: 433-444.
- Kim, Mi-Ah, Jae-kwan Lee, Kyung-Duk Zoh (2007) Evaluation of the geum river by multivariate analysis: principal component analysis and factor analysis. J. Korean society on Water Quality. 23(1): 161-16.
- Kim, B. H., S. E., Lee, M. Y., Song, J. H., Choi, S. M., Ahn, K. S., Lee, E.

- C., Cho, T. S., Chon, S. C., Koh (2008) Implementation of artificial neural networks (ANNs) to analysis of inter-taxa communities of benthic microorganisms and macroinvertebrates in a polluted stream. *Science of the Total Environment*. 390: 262-274.
- Klemm, D. J., K. A., Blocksom, F. A., Fulk, A. T., Herlihy, R. M., Hughes, P. R., Kaufmann, D. V., Peck, J. L., Stoddard, W. T., Thoeny, M. B., Griffith and W. S. Davis (2003) Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing Mid-Atlantic Highlands streams. *Environmental Management*. 31(5): 656-669
- Kohonen, T. (1984) *Self-Organization and Associative Memory*. Springer-Verlag. Berlin.
- Lafontaine, Y., F. Gagne, C. Blaise, G. Costan, P. Gagnon and H. M. Chan (2000) Biomarkers in zebra mussels (*Dreissena polymorpha*) for the assessment and monitoring of water quality of the St Lawrence River (Canada). *Aquatic Toxicology*. 50: 51-71.
- Legendre, P. and L. Legendre (1983) *Numerical Ecology*. Elsevier Scientific Publishing Company, Netherlands. p. 419
- Lippmann, R. P. (1987) An Introduction to Computing with Neural Nets. *IEEE ASSP Magazine*, April. p. 4 - 22
- Melsa, J. L., and D. L., Cohn (1978) *Decision and estimation theory*. New York : McGraw-Hill.
- Muyzer, G, E. C., Waal and A. G. Uitterlinden (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl and Envir. Microbiol.*, 59: 695-700.
- Merritt, R. W. and K. W., Cummins (1984) *An Introduction to the Aquatic Insects of North America*. Hunt Publishing Company, Dubugue.
- National Water Council (1981) *River Quality: The 1980 Survey and Future Out-look*. London, National Water Council.
- Pennak, R. W. (1978) *Fresh-water Invertebrates of the United States*. John

- Wiley & Sons, Inc., New York. p. 803
- Petersen, W., L., Bertino, U., Callies and E., Zorita (2001) Process identification by principal component analysis of river water-quality data. *Ecological Modeling*. 138: 203-213.
- Rapport, D. J., H. A. Regier and T. C. Hutchinson (1985) Ecosystem behavior under stress. *Am. Naturalist*. 125: 617-640.
- Rast, W. and M. Holland (1988) Eutrophication of lake and reservoirs: a framework for making management decisions. *Ambio*. 17: 2-12.
- Resh, V.H. and D.M. Rosenberg (1984) *The ecology of aquatic insects*. Praeger publishers.
- Rockhow, K. H. (1999) Water quality prediction and probability network models. *Can. J. Fish. Aquat. Sci.* 56: 1150-1158.
- Roth, N., M. Southerland, J. Chaillou, R. Klauda, P. Kazyak, S. Stranko, S. Weisberg, L. Hall, R. Morgan Ii (1998) Maryland Biological Stream Survey: Development of a Fish Index of Biotic Integrity. *Environmental Monitoring and Assessment*. 51(1): 89-106.
- Schaeffer, D. J., E. E. Herricks and H. W. Kerster (1988) Ecosystem health measuring ecosystem health. *Environment Manage.* 12: 445-455.
- Shapiro, J., V., Lamarra and M., Lynch (1975) Biomanipulation: an ecosystem approach to lake restoration. In Brezonik, P. L. and J. L., Fox(eds). *Proc. Sym. Wat. Qua. Man. Bio. Con., Univ. Fla. Gainesville*: 85-96.
- Simon, T. P. and J. Lyons (2005) *Application of the Index of Biotic Integrity to Evaluate Water Resource Integrity in Freshwater Ecosystems*: Davis, W. S. and T. P. Simon. *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. CRC Press. Boca Raton.
- Speckt, D. F. (1990) Probabilistic neural networks. *Neural Networks*. 3: 109 - 118.
- Thierfelder, T. (1999) *Standard Probability Density Functions of Characterizing Water Quality Variables in Dimictic Glacial/Boreal*

- Lakes. *Water, Air, and Soil Pollution*. 113: 295-317.
- U.S.EPA (1985) Technical support document for water quality-based toxic control. Office of water Enforcement and Permits, Office of Regulations and Standards. Washington, D.C., USA.
- U.S.EPA (1991) Technical support document for water quality-based toxic control. EPA 505-2-90-001. Office of water, Washington, D.C., USA.
- U.S.EPA (1996) Guideline for Preparation of the State Water Quality Assessments (305(b) Report).
- Wang, X. (2001). "Integrating water-quality management and land-use planning in a watershed context." *J. Environmental Management* 61(1): 25-36.
- Watanabe T, K. Asai and A. Houki (1986) Numerical estimation of organic pollution of flowing water by using the epilithic diatom assemblage: Diatom Assemblage Index (DAIpo). *Sci. tot. Environ.* 55: 209-218.
- Weigel, B. M., L. J. Henn, E. M. Martinex-rivera (2002) Macroinvertebrate based index of biotic integrity for protection of streams in west-central Mexico. *J. N. Am. Benthol. Soc.* 21(4): 686-700
- Wolf, L. and S. Bileschi (2005) Combining variable selection with dimensionality reduction. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2: 801-806.
- Xu, F. L., S. T, R. W. Dawson, P. G. Li and J. Cao (2001) Lake ecosystem health assessment: Indicators and methods. *Wat. Res.*, 35: 3157-3167
- Xu, F. L., Z. Y. Zhaoa, W. Zhan, S. S. Zhao, R. W. Dawson, S. Tao (2005) An ecosystem health index methodology (EHIM) for lake ecosystem health assessment. *Ecological Modelling*. Online.