

Research Paper

지표의 내재적 가중치를 고려한 하천의 생태적 건전성 평가

김홍명 · 하성룡

충북대학교

Analyzing Ecological Soundness Considering the Implicit Weight of the Indicator

Hong-Myung Kim · Sung-Ryong Ha

Department of Urban Engineering, Chungbuk National University

요약: 본 연구의 목적은 금강수계 하천의 생태적 건전성을 평가하는 체계를 구축하는 것이다. 연구대상 지역은 금강수계 14개 중권역으로 하였다. 평가지표의 투명성과 일관성을 확보하기 위하여, 건전성 지표 선정은 지표 간 내재적 가중치 변화 특성을 고려한 조정방법을 사용하여 선정하였다.

최종 통합지수에 미치는 영향이 가장 큰 지표는 수량, 수질, 수생태, 서식·수변환경 부문 중 수생태 부문의 지수로 파악되었다. 하천의 생태적 건전성 지수를 분석한 결과, 댐상류 유역(대청댐 기준)은 2014년까지 기준년도(2008)에 비해 상대적으로 좋은 상태로 평가되었으며 댐하류 유역은 좋지 않은 상태로 평가되었다. 년 단위 생태적 건전성 지수의 변화 추이는 다음과 같다. 용담댐, 용담댐하류, 보청천, 대청댐, 대청댐하류, 논산천의 경우는 시기별 차이는 있으나 건전성 지수의 저하가 진행되고 있다. 반면, 무주남대천, 영동천, 갑천의 경우는 건전성이 개선되고 있는 것으로 평가되었으며 초강, 대청댐상류, 미호천, 금강공주, 금강하구연은 건전성이 개선된 후 다시 저하되는 것으로 평가되었다.

주요어: 하천, 생태적 건전성 지수, 내재적 가중치, 가중치의 중요도, 통합지표 집계

Abstract : The purpose of this study is to establish a system to evaluate the ecological soundness of the Geum river basin. The study target area is 14 sub-watersheds of the Geum river basin. For the selection of indicators to ensure transparency and consistency of the evaluation indicators, the ecological soundness indicators were secured by using the indicator adjustment method derived in consideration of the intrinsic weight change characteristics between indicators.

The index with the greatest impact on the final composite index was identified as the index of the aquatic ecology among the water quantity, water quality, aquatic ecology, and habitat-riparian environment dimensions. As a result of analyzing the ecological health index of the river, the watershed upstream of the dam (based on the Daecheong -dam) was evaluated to be in relatively good condition until 2014 compared to the base year(2008), and the watershed downstream of the

dam was evaluated to be in a poor condition.

The annual trend of changes in the ecological soundness index on an annual basis is as follows. In the case of Yongdamdam, Yongdamdamdownstream, Bocheong-chun, Daechungdam, Daechungdamdownstream, and Nonsancheon, although there are differences by time period, the soundness index is in declining. On the other hand, Mujunamdaecheon, Yeongdongcheon, and Gapcheon were evaluated to have improved soundness, while Chogang, Daechungdamupstream, Mihocheon, Gongjugeumgang, and Geumgangestuary were evaluated to deteriorate again after soundness was improved.

Keywords : River, Ecological soundness index, Implicit weight, Importance of the weight, Aggregation

I. 서론

하천은 다양한 생물 종이 안전하고 풍족한 서식 및 군락 공간으로 생태계 기능은 물론이고 인간의 생활, 농업, 농업에 필요한 물 공급 및 교통, 자연경관의 청명함을 경험할 수 있는 관광 등 사회적인 기능도 공유하는 공간이 되어준다. Pinto et al. (2011)은 하천 생태계의 이해는 생태계와 인간사회 간의 이해충돌을 해소하는 기본이라 했다. 하천 생태계의 건강성 평가에는 다양한 평가지표가 필요하고 특히, 지표 간의 관계를 잘 이해하고 정량적 지표 선정을 위해서 선정 기준은 물론이고 그 적용을 위한 연구자의 책임이 요구된다.

Rowley et al. (2012)는 종합지수 집계(aggregation)에 사용되는 가중치(weight)의 할당에는 연구자의 책임과 역할이 중요함을 강조했다. Munda & Nardo (2005)는 지표의 가중치 할당은 지표 평가체계에서 개체지표 간의 상대적 중요성을 반영하기 위해 사용되며 또한 이러한 명시적(explicit) 가중치를 집계에 적용함으로써 종합지수가 산출됨을 제시했다.

한편, Paruolo et al. (2013), Becker et al. (2017)은 지표 집계 과정에 있어 지표의 비선형성과 데이터 구조 또한 지표 간의 암묵적 상관관계 등으로 산출된 통합지수 값이 할당된 명시적 가중치를 제대로 반영하지 못한다는 연구결과를 제시했다. 즉, 지표 간의 상대적 중요성은 명시적으로 할당된 가중치 체계와 일치하지 않을 수 있으며 지표의 집계 과정에서 각 변수의 서로 암묵적 의존 관계로 인하여 할당된 가중치가 변화될 수도 있음을 강조했다. 그리고 지표 간

의 암묵적인 의존 현상을 ‘내재적 가중치(implicit weight)’라 정의했다. 내재적 가중치는 결국 변수의 특성에 의해 좌우되며 이러한 변수 간에 내재되어 있는 특성을 반영하여 종합지수를 산정할 수 있다면 개발되는 종합지수의 설명력 향상과 신뢰성 확보를 기대할 수 있다. 따라서 본 연구는 하천 생태계 자료의 내재적 가중값을 고려한 금강수계 중권역 수준의 하천이 가지는 생태적 건전성 평가체계를 구축함을 목적으로 한다.

II. 연구내용 및 방법

1. 연구대상 유역

금강수계는 금강 최상류 용담댐 유역, 대청댐 유역 그리고 대청댐 하류 유역으로 구분된다. 하천 건전성 평가를 위한 대상 지역은 금강수계를 14개 권역으로 분할한 ‘중권역’을 대상으로 하였다(Figure 1). 지역 단위로는 대청댐을 기준으로 댐 상류 및 댐 하류 지역으로 구분 평가했다. 댐 상류 지역에 해당하는 중권역은 용담댐(3001), 용담댐하류(3002), 무주남대천(3003), 영동천(3004), 초강(3005), 대청댐상류(3006), 보청천(3007), 대청댐(3008)으로 총 8개 중권역이다. 댐하류 지역에 해당하는 중권역은 갑천(3009), 대청댐하류(3010), 미호천(3011), 금강공주(3012), 논산천(3013), 금강하구연(3014)으로 총 6개 중권역이다. 댐 하류지역 중 갑천, 미호천, 논산천 권역은 댐에 의해 직접적인 영향을 받지 않는 유역으로 구분된다.

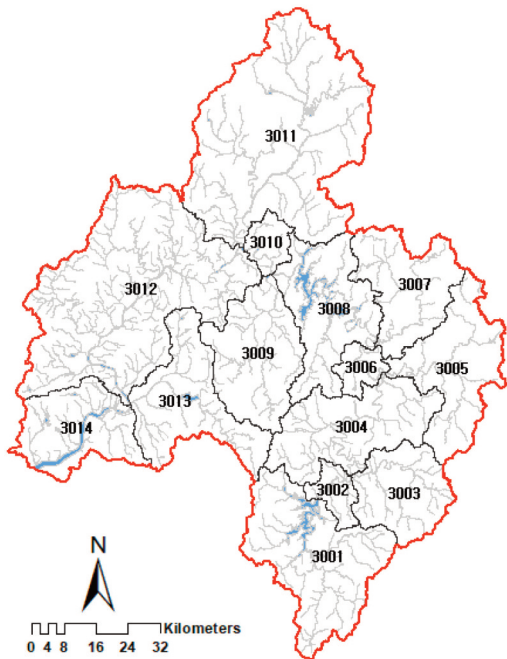


Figure 1. Study area: Geumgang river watershed.

2. 지표의 구축 및 정규화

1) 지표의 구축

하천의 생태적 건전성 지수(ESI, ecological soundness index) 구축을 위한 지표 체계는 고유량, 저유량, 영유량, 변동성, 계절성 등 수량 부문 지표 5개, BOD, 총인(TP), 총대장균군(TC), 클로로필-a (Chl-a) 그리고 BOD와 TP의 목표수질 달성도 등 수질 부문 6개 지표, 부착조류(TDI), 저서성무척추대형동물(BMI), 어류(FAI) 등 수생태 부문 3개 지표, 서식 및 수변환경(HRI)의 서식·수변 부문 1개 지표로 총 4개 부문, 15개 지표로 구성하였다. 기초자료 구축에 필요한 자료는 한국수자원공사 (2018)에서 수행한 하천 종합건전성 평가체계 연구 자료를 인용했다.

2) 지표의 정규화

지표의 정규화는 기준년에 따른 지수의 절대적인 비교 및 해석을 위해 Mazziotta et al. (2016, 2018)이 제시한 Min-Max 재조정 방법을 적용하였다. 주어진 매트릭스 $X = \{x_{ij}\}$ 의 정규화는 식 (1)을 사용하여 매트릭스 $R = \{r_{ij}\}$ 를 계산한다.

$$r_{ij} = \frac{(x_{ij} - Min_x)}{(Max_x - Min_x)} 60 + 70 \quad (1)$$

여기서, Min_x 와 Max_x 는 지표 j 의 최대와 최소값이다(이하 “goalposts”라 한다). 지표 j 가 음(-)의 값을 가지면 식 (1)은 200으로 보완하여 계산한다. 용이한 결과 해석을 위해 Goalposts는 100으로 대표되는 기준값(예: 기준 연도의 평균)으로 고정할 수 있고 Goalposts를 설정하기 위한 간단 절차는 다음과 같다. 모든 단위와 시간을 고려한 지표 j 의 전체 최솟값과 최댓값을 Inf_x 와 Sup_x , 지표 j 에 대한 기준값을 Ref_x 하면 ‘goalposts’는 아래와 같이 정의된다.

$$\begin{cases} Min_x = Ref_x - \Delta \\ Max_x = Ref_x + \Delta \end{cases} \quad (2)$$

여기서 $\Delta = (Sup_x - Inf_x)/2$, 정규화된 값이 100이면, 기준값은 70~130의 범위를 갖는다.

3. 평가 지표 선정

1) 변수와 가중치의 중요도 측정

(1) 종합지수에 대한 변수의 중요도 측정

일반적으로 종합지수는 Figure 2와 같이 지표들에 의해 잠재변수(latent)가 영향을 받는 조형 측정 모델(formative measurement model)이며 또한 지표 간의 상관관계(Φ_{ij})에 의한 암묵적인 가중치가 존재한다(Diamantopoulos et al. 2008). 지표들의 중요도 수준을 나타내기 위하여 할당된 명시적 가중치, w 는 내재적 가중치에 의하여 실제 지표들의 중요도에 반하는 결과를 가져온다. 따라서 지표들의 집계 과정에서

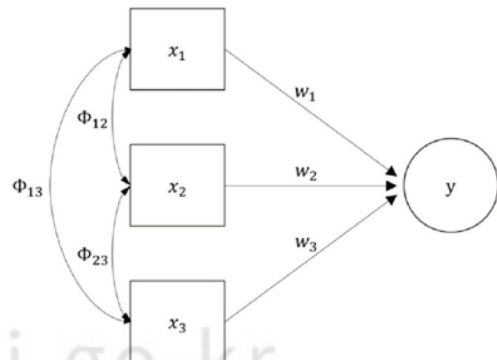


Figure 2. Formative measurement model.

변수 및 가중치의 중요도 측정은 매우 중요하다.

(2) 비선형 상관비, S_i

변수 및 가중치의 중요도는 종합지수에 대해 각 변수가 갖는 영향을 조건부 기대 분산을 이용한 비선형 상관비, S_i 를 측정함으로써 산출할 수 있다. 종합지수에 대한 각 입력 변수 x_i 의 영향을 평가하는 방법은 x_i 에 대한 y 의 의존성을 측정하는 것이다. 여기서 변수는 정규화된 것으로 가정한다.

$$y_i = f_i(x_{ji}) + \epsilon_j \quad (3)$$

여기서 $f_i(x_{ji})$ 는 샘플 포인트 x_{ji} 가 주어진 경우, y 의 조건부 평균을 모형화하는 적정 함수이며 ϵ_j 는 x_i 이외의 지표로 인한 변동을 설명하는 오차 항이다. x_i 에 대한 y 의 의존성은 x_i 와 y 사이의 공분산과 상관계수로 식 (4)와 같이 정의한다.

$$\begin{aligned} cov(y, x_i) &:= E[(y - \mu_y)(x_i - \mu_x)], \\ R_i &:= corr(y, x_i) := \frac{cov(y, x_i)}{\sigma_y \sigma_x} \end{aligned} \quad (4)$$

여기서, $cov(y, x_i)$ 는 구간 $[-1, 1]$ 을 가진 피어슨 적률 상관계수(Pearson product moment correlation coefficient)라고 불리는 공분산의 표준화이다. 단순 선형 회귀 분석의 경우, 상관계수의 제곱으로 결정 계수 R_i^2 , 즉, 구간 $[0, 1]$ 의 값을 갖는 $R_i^2 = corr^2(y, x_i)$ 이다.

R_i^2 이 선형 의존성을 측정하는 것처럼 결정 계수는 일반 선형 회귀의 적합도(goodness of fit)를 측정하는 데 사용된다. 식 (4)의 공분산, 상관계수 및 결정 계수 R_i^2 은 모든 선형 의존성과 관련된 측정값이다. 예로, 결정 계수 R_i^2 은 식 (5)로 계산된다.

$$SS_{reg,i} / SS \quad (5)$$

여기서, $SS_{reg,i} = \sum_{j=1}^n (\hat{f}_i(x_{ij}) - \bar{y})^2$ 는 선형 회귀로 설명되는 제곱의 합(변동)이며 $\bar{y} := n^{-1} \sum_{j=1}^n y_j$ 는 표본 평균, $\hat{f}_i(x_{ij}) = \hat{\beta}_0 + \hat{\beta}_1 x_{ij}$ 는 관측되는 y_j 에 대한 선형 회귀 적합 곡선, $SS = \sum_{j=1}^n (y_j - \bar{y})^2$ 는 제곱의 총합이다.

따라서 R^2 은 x_i 에 대한 y 의 선형 회귀에 의해 설명된 제곱 합과 y 의 총 제곱 합의 비율로 볼 수 있다. 이 측정은 선형 회귀를 기반으로 y 와 x_i 사이의 비선형성

을 설명하지 않는다.

종합지수에서 집계 공식은 종종 선형이지만 변수 사이의 상관관계에 의해 x_i 에 대한 y 의 관계가 비선형일 수 있다. 이런 경우, R_i^2 은 의존성의 정도를 과소 추정할 수 있다. 여기에서 채택된 비선형 효과인 상관비 $S_i, i = 1, 2, \dots, d$ 는 1차 감도지수 또는 주 영향 지수로 널리 알려져 있다(Becker et al. 2017). 이 측정은 종합지수에 대한 각 변수의 비선형 영향을 측정하기 위한 것이며 $f_i(x_{ij})$ 가 비선형일 때 선형인 R_i^2 이 S_i 가 되는 비선형 일반화이다. 만약, 주어진 변수가 고정되어 있다면 종합지수 값에 대한 기대 분산 감소(expected variance reduction)로 해석될 수 있으며 식 (6)으로 정의한다.

$$S_i \equiv \eta_i^2 := \frac{V_{x_i}(E_{X_{-i}}(y | x_i))}{V(y)} \quad (6)$$

여기서 X_{-i} 는 변수 x_i 를 제외한 모든 변수(x_1, \dots, x_d)를 포함하는 벡터이다. $E_{X_{-i}}(y | x_i)$ 는 첨자 X_{-i} 와 함께 명시적으로 나타내며 x_i 의 주어진 값에서 y 의 기댓값은 X_{-i} 에 대한 기댓값이다. 첨자 X_{-i} 는 개념적 혼란을 막기 위해 생략하고 $E_{x_i}(y | x_i) \equiv E(y | x_i)$ 로 나타낸다. 즉, x_i 가 변동 구간에서 하나의 값으로 고정되었을 때 x_i 에 대한 조건부이다. $E(y | x_i)$ 는 x_i 의 주 효과(main effect)로 알려져 있으며 앞서 설명한 $f_i(x_{ji})$ 와 동등하다. 따라서 $E(y | x_i)$ 는 x_i 에 대한 y 의 산포도(Scatter plot)에 대한 비선형 회귀 적합에 지나지 않는다.

j 번째 샘플 포인트 y_j 와 일치하는 회귀 적합 값(x_i 에 대한 y)을 $m_{ji} := \hat{f}_i(x_{ji})$ 라고 하면 상관비 S_i 는 아래 식 (7)과 같이 추정할 수 있다. 여기서, \bar{m}_i 는 m_{ji} 의 평균이다.

$$\hat{S}_i = \frac{\sum_{j=1}^n (m_{ji} - \bar{m}_i)^2}{\sum_{j=1}^n (y_j - \bar{y})^2} \quad (7)$$

결과적으로, x_i 에 대한 y 의 비선형 의존도인 상관 비율을 x_i 에 대한 y 의 산포도에 비선형 회귀를 적합시키고 결과 곡선의 분산을 취하여 y 의 무조건부 분산으로 표준화하면 간단히 추정할 수 있다(Becker et al. 2017).

(3) 집계로부터의 분리

입력 매개 변수가 상호 연관될 때 상관 비율 S_i 는

Table 1. Criteria for Selecting Indicators by EFA

Item	Criteria	Explanation
KMO measure of sampling adequacy	More than 0.5	Identify common factors suitable for analysis
Bartlett's test of sphericity	Not more than p-value 0.1	
MSA	More than 0.5	Evaluate whether individual variables should be included in factor analysis
Communality (CM)	More than 0.4	Explanation ability of variable variables for each factors
Eigen value	More than 1	Explanation ability of more than 1 variable.

각 변수 x_i 의 분산 기여도를 포착할 뿐만 아니라 x_i 와 관계가 있는 다른 변수들에 의한 기여도로 분석할 수 있다(Xu & Gertner 2008; Mara & Tarantola 2012). 따라서 각 변수의 영향 특성에 대한 통찰력을 얻기 위해 S_i 를 다음과 같이 두 부분으로 분해할 수 있다.

$$S_i = S_i^c + S_i^u \quad (8)$$

여기서, S_i^c 는 상관 기여(correlated contribution)로 다른 지표와 상관관계가 있는 지표의 변화량을 나타내고 S_i^u 는 비상관 기여(uncorrelated contribution)로 다른 지표로는 설명할 수 없는 지표 자체의 변화량을 나타낸다. 이 구분을 통해 각 지표의 영향을 상관 기여 및 비상관 기여로 나누고 개발자에게 할당된 가중치가 종합지수 산정 시 각 지표의 대표성에 어떤 영향을 갖는지에 대한 추가 이해를 제공한다(Xu & Gertner 2008; Becker et al. 2017).

2) 지표 선정

지표를 선정하는 방법으로 기존에 많이 사용된 탐색적 요인분석(exploratory factor analysis: EFA)을 통하여 1차 지표를 선정하였다. 1차 선정된 지표를 다시 상관관계 분석 및 각 지표가 종합지수에 갖는 비선형 상관비(non-linear correlation ratio) 분석을 통해 지표를 재조정하여 최종 지표를 선정하였다.

(1) 탐색적 요인분석에 의한 지표 선정

구축된 지표에 대해 1차 지표 선정을 위한 주요 기준항목으로 표본 적합도 검사(Kaiser-Meyer-Olkin: KMO), Bartlett 단위행렬 검정(Bartlett's test of sphericity), 표본 적합성 측도(measure of sampling adequacy: MSA), 공통성(communality)

및 고유치(eigen value)를 사용했다(Table 1). KMO는 잔영상관행렬(ant-image correlation matrix)을 이용하여 계산된 통계량으로 자료가 요인분석에 적합한지를 판단하며 MSA는 개별 변수를 요인분석에 포함해야 하는 여부를 평가한다(Kim 2016; Hauben et al. 2017; Joo et al. 2018).

(2) 상관관계 및 비선형 상관비에 의한 지표 선정

상관관계 분석에서 일반적으로 상관계수 r 값이 0.8 이하 또는 0.3 이상일 때 개발하고자 하는 종합지수 산정에 적합한 지표로 볼 수 있다. 상관계수 값이 낮거나 음의 값을 가지는 경우 종합지수에 미치는 영향이 미미하거나 역효과를 유발하는 경향이 있다. 그러나 상관계수 값이 낮거나 음의 값을 가져 종합지수 구성에 미치는 영향이 미미하더라도 지표 자체가 가지는 특성을 무시할 수는 없으며 이를 비 대체성(non-substitution)이라 한다(OECD 2008). 따라서 상관계수 값이 실제 음(-)의 값을 나타내더라도 주 효과를 나타내는 비선형 상관비 분석을 통해 각 지표가 지수에 갖는 실제 영향력은 양(+)의 영향을 나타낼 수 있다. 즉, 종합지수 산정 시 이러한 지표의 비 대체성을 고려한 집계가 되어야 한다. 반면 값이 음(-)인 부의 영향을 나타내는 지표는 집계에서 제외하였다(Lindén 2018).

상관계수가 매우 큰 값을 가지면 해당 지표들의 공선성(collinearity) 문제로 인하여 종합지수 집계 시 이중 계산(double count)되어 보상성(compensability)의 문제를 유발한다(Kao et al. 2008; Saisana & Saltelli 2011). 따라서 문제를 제거하기 위한, 새로운 대체지표를 설정한 다음 상관계수가 높은 해당 지표들을 대체지표의 하위 부문(sub-dimension)으로

두고 산술평균한 후에 평균값을 대체지표 값으로 반영한 후 집계에 포함하였다.

(Variability) 지표의 MSA 값은 각각 0.579, 0.611, 0.706으로 양호한 것으로 판정했다. 이로써 저유량, 영유량 및 변동성 3개를 1차 지표로 선정하였다.

III. 결과 및 고찰

1. 지표 선정 결과 분석

1) 수량 부문

(1) 탐색적 요인분석에 의한 지표 선정

KMO는 0.500, Bartlett 검정은 근사 카이제곱 520.859, 유의확률 $p = 0.000$ 으로 적합한 것으로 나타났다. Table 2에서 보듯이 요인분석에 의한 지표의 신뢰성 분석결과, 계절성(Seasonality), 고유량(High flow) 지표의 MSA 값이 각각 0.348 와 0.271로 0.5 이하로 지표 적합성에 문제가 있는 것으로 판단하고 지표에서 제외하였다.

부적합 지표를 제외한 후 재차 실시한 요인분석 결과, 저유량(Low flow), 영유량(Zero flow), 변동성

(2) 상관관계 및 상관비 S_i 에 의한 지표 선정

각 지표가 지수에 미치는 영향(또는 비선형 의존성)을 추정하는 집계방식은 산술평균을 사용하고 가중치 배정은 모든 지표에 동일하게 적용했다. 비선형 상관비 S_i 산정은 비선형 회귀 적합한 벌칙 스플라인 곡선(penalized splines)을 사용했다(Figure 3). 여기서 벌칙 스플라인 곡선은 모든 지표를 포함하는 종합지수 값과 각 지표 x_i 를 고정(제외)할 때, 나머지 지표들에 의해 산출된 지수 값의 적합 곡선을 의미한다. 본 연구에서 상관비 S_i 추정은 Becker (2017)의 방법을 인용하여 Matlab 프로그램을 사용했다.

지표 간 상관관계 분석결과, 저유량, 영유량, 변동성 지표의 계수 값은 0.3119~0.5857로 나타났으며 OECD (2008) 기준의 상관계수($|r| < 0.8$ 또는 $|r| >$

Table 2. Results of EFA in Water quantity

Indicator	Factor loading	Communality	MSA	Eigen value	Accumulated variance (%)
Low flow	0.871	0.964	0.570	1.914	38.1
Zero flow	0.812	0.762	0.567		
Variability	0.699	0.745	0.715		
Seasonality	0.970	0.571	0.348	1.073	59.0
High flow	0.982	0.943	0.271	0.998	76.6

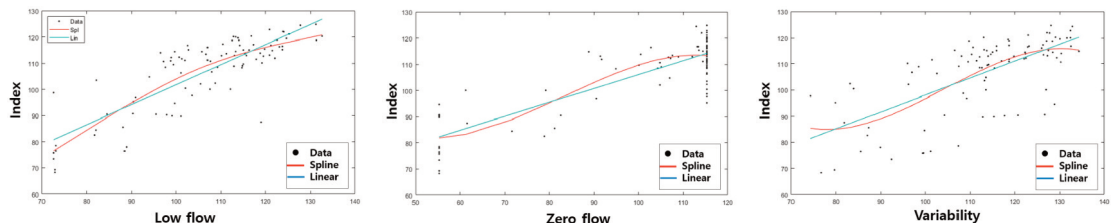


Figure 3. Spline regression fits of the index for each indicator in water quantity.

Table 3. Results of calculation in water quantity

Indicator	S_i^u	S_i^c	S_i	CR* (%)
Low flow	0.0000	0.7139	0.7139	37.1
Zero flow	0.0009	0.7391	0.7400	38.4
Variability	0.1885	0.2822	0.4707	24.5

*: Contribution rate

0.3)을 만족하였다.

Figure 3은 입력 지표들의 산포도($E(y | x_i)$)와 주 효과의 단순 선형 및 비선형 벌칙 스플라인 추정을 나타낸 것으로 단순 선형 및 비선형 벌칙 스플라인 적합은 서로 상당히 비슷했다. Table 3과 같이 상관비 S_i 는 비상관 기여(S_i^0)과 상관 기여(S_i^1)로 구분할 수 있으며 비상관 기여 S_i^0 는 해당 지표가 다른 지표와의 상관관계로 인한 영향을 적게 받으며 지표 자체가 가진 고유한 변화량을 의미하며 상관 기여가 S_i^1 값에 가까워질수록 가치치 값이 최종 결과에 거의 영향을 미치지 않을 수 있음을 의미한다. 또한, 지표 간 부(-)의 상관관계로 인하여 상관 기여가 음수이면 변수에 개별적 문제가 있음을 나타낸다. 따라서 집계에서 제외하는 것이 바람직하다(Becker et al. 2017).

수량 부문 지표의 상관비 값을 계산한 결과, 저유량 37.1%, 영유량 38.4%, 변동성 24.5%로 변동성 지표가 상대적으로 지수에 미치는 영향이 낮은 것으로 분석되었다. 또한, 변동성 지표는 비상관 기여가 약 40% 정도의 기여율을 갖고 있어 지표 자체가 독립적인 특성을 갖는 것으로 나타났다.

2) 수질 부문

(1) 탐색적 요인분석에 의한 지표 선정

KMO 값은 0.764, Bartlett 검정은 근사 카이제곱 594.339, 유의확률 $p = 0.000$ 으로 선정 기준을 만족하는 것으로 나타났다. 요인분석에 의한 지표의 신뢰성 분석결과, 공통성 및 고유치(4.407) 또한 모두 만족하는 것으로 분석되었다. MSA 값은 0.720~0.843의 범위로 적합하였으며 누적분산량도 73.4%로 적절하였다(Table 4).

(2) 상관관계 및 상관비에 의한 지표 선정

Figure 4는 입력 지표들에 대한 지수의 산포도로 선형 및 비선형 벌칙 스플라인의 적합 곡선을 살펴보면 BOD, TP, BOD의 목표수질 달성도(T-BOD)는 상당히 비슷하였으며 총대장균군, 클로로필-a(Chl-a), TP의 목표수질 달성도(T-TP)는 명확한 차이를 나타냈다.

지표의 상관관계 분석결과, BOD는 TP(0.908) 총대장균군(0.817), 클로로필-a(0.811) 지표와 유의수준 $p < 0.01$ 에서 높은 수준($|r| \geq 0.8$)의 상관관계를 나타냈다. BOD 지표는 높은 상관관계로 집계 시 공선성 문제를 유발하여 이중 계산의 문제가 발생할 수 있다. 따라서 본 연구에서는 해당 지표들을 하위 차원(dimension)으로 두고 상관계수가 큰 값들을 순서대로 평균하여 새로운 지표로 변환한 후 상관관계를 재조정된 결과, BOD, TN, 총대장균군(TC)의 조정 상관계수 값은 지표 간의 공선성 문제가 유발하지 않는 수준에 부합하였다.

새롭게 선정된 지표를 “유기물(OM, organic matter)”로 명명하여 상관관계를 재분석한 결과, 유기물, 클로로필-a, 목표수질 달성도(T-BOD, T-TP) 지표의 계수 값은 0.4340~0.7166으로 지표 선정 기준에 부합했다(Table 5). S_i 분석결과, 각 지표가 종합지수에 미치는 영향은 거의 대등하였다. 목표수질 달성도(T-BOD, T-TP)는 집계 시 다른 지표에 의해 영향을 받는 상관 기여가 높게 나타났으며 유기물, 클로로필-a는 비상관 기여도 대등한 영향을 가진 것으로 나타났다(Table 6).

Table 4. Results of EFA in water quality

Indicator	Factor loading	Community	MSA	Eigen value	Accumulated variance (%)
BOD	0.950	0.902	0.727	4.407	73.4
TP	0.945	0.892	0.843		
TC	0.815	0.664	0.733		
Chl-a	0.798	0.636	0.720		
T-BOD	0.860	0.740	0.830		
T-TP	0.756	0.572	0.726		

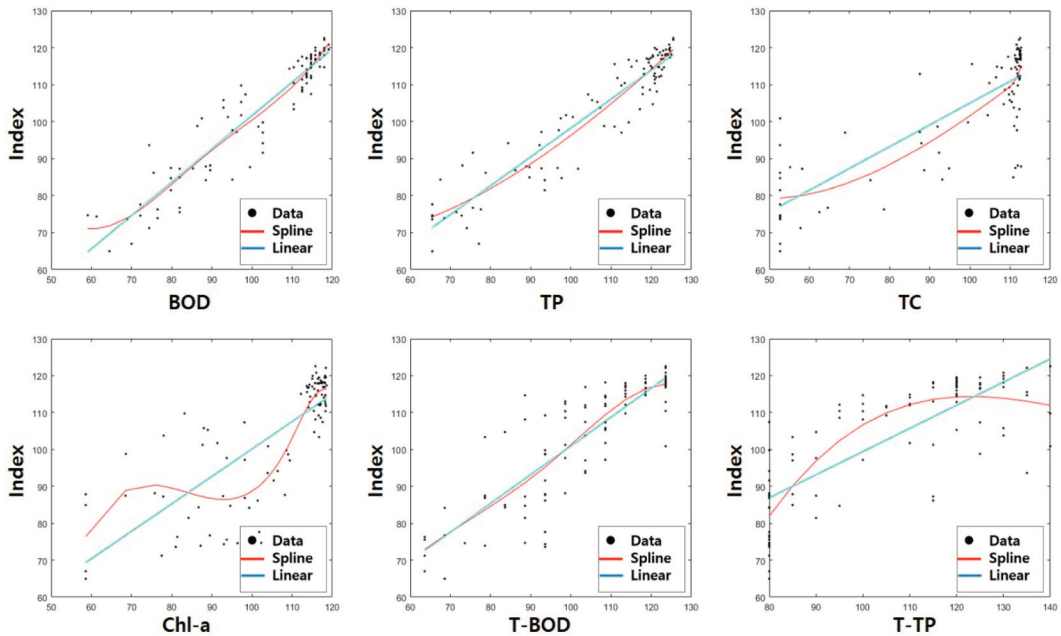


Figure 4. Spline regression fits of the index for each indicator in water quality.

Table 5. Results of re-correlation analysis in water quality

Indicator	OM*	Chl-a	T-BOD	T-TP
OM	1	0.6913	0.7166	0.6171
Chl-a	0.6913	1	0.6395	0.4340
T-BOD	0.7166	0.6395	1	0.7381
T-TP	0.6171	0.4340	0.7381	1

*: Organic matter

Table 6. Results of calculation in water quality

Indicator	S_i^u	S_i^c	S_i	CR (%)**
OM*	0.3276	0.4824	0.8100	26.3
Chl-a	0.3467	0.3573	0.7040	22.9
T-BOD	0.0000	0.8158	0.8158	26.5
T-TP	0.0000	0.7509	0.7509	24.4

*: Organic matter, **: Contribution rate

3) 수생태 및 서식·수변 부문

(1) 수생태 부문

수생태 부문의 지표는 요인분석 및 상관관계, 주영향 분석에서 모두 지표의 적합성에 만족하는 것으로 나타나 모든 지표를 집계에 포함했다. 분석결과를 살펴보면, KMO 값은 0.707로 나타났으며 Bartlett 검정은 근사 카이제곱 127.770, 유의확률 $p = 0.000$,

고유치 2.304로 지표 선정 기준에 부합하는 것으로 나타났다. 공통성과 MSA 값은 부착돌말류(TDI, trophic diatom index), 저서성무척추동물(BMI, benthic macro-invertebrate index), 어류(FAI, fish assessment index)가 각각 0.707, 0.771, 0.826 과 0.784, 0.704, 0.659로 모두 지표의 적합성에 만족하는 것으로 나타났다.

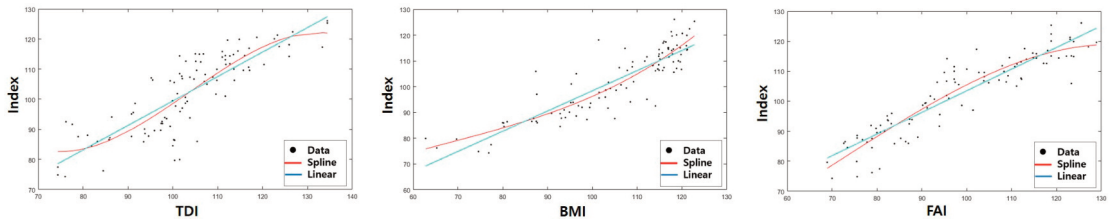


Figure 5. Spline regression fits of the index for each indicator in aquatic ecology.

Table 7. Results of calculation in aquatic ecology

Indicator	S_i^a	S_i^c	S_i	CR* (%)
TDI	0.1586	0.5410	0.6996	30.0
BMI	0.0001	0.7840	0.7841	33.6
FAI	0.1953	0.6550	0.8503	36.4

*: Contribution rate

상관관계 분석결과, 각 지표 간 상관계수의 범위는 0.5758~0.7278로 지표 선정 기준을 만족하는 것으로 나타났다. Figure 5는 입력 지표들에 대한 산포도로 선형 및 비선형 벌칙 스플라인의 적합 곡선으로 3개 지표 모두 상당히 비슷한 것으로 나타났다.

상관비 S_i 분석결과, 각 부착돌말류, 저서성무척추 대형동물, 어류 지표가 종합지수에 미치는 영향 (CR, %)은 30.0, 33.6, 36.4로 큰 차이가 없는 것으로 나타나 수생태 부문의 모든 지표는 배제 없이 종합 집

계에 포함했다. 그러나 부착돌말류와 어류 지표는 지표 자체가 종합지수에 대해 일정량의 독립적인 영향을 갖는 비상관 기여도 있는 것으로 나타났다(Table 7).

한편, 서식·수변 부문은 서식및수변환경(HRI, habitat riparian index) 지표가 1개로 별도의 분석은 수행하지 않고 종합 집계에 포함했다.

4) 최종 지표 선정 결과

Table 8은 지표 조정과정을 통하여 결정된 지표의

Table 8. Analysis results before and after selection of indicators

Dimension	Indicator alteration		Alteration reason
	Before	After	
Water quantity	High flow	Exception of aggregation	Low anti-image correlation coefficient (0.348)
	Low flow	Low flow	
	Zero flow	Zero flow	
	Variability	Variability	
	Seasonality	Exception of aggregation	Low anti-image correlation coefficient (0.271)
Water quality	BOD	Organic matter	Double counting by high correlation in inter variables
	TP		
	TC		
	Chl-a	Chl-a	
	T-BOD	T-BOD	
	T-TP	T-TP	
Aquatic ecosystem	TDI	TDI	
	BMI	BMI	
	FAI	FAI	
Habitat-riparian environment	HRI	HRI	

전과 후의 결과 비교이다. 탐색적 요인분석, 상관관계 및 S_i 분석에 따른 지표 선정결과, 수량 부문의 고유량, 계절성 지표는 표본 적합성 측도인 MSA 값이 낮게 측정되어 지표를 배제하였다.

높은 상관관계를 가진 지표는 집계 시 공선성의 문제를 유발하나 지표를 제외하기보다는 해당 지표들을 평균하여 새로운 지표로 집계에 포함했다. 높은 상관관계 값을 갖는 BOD, TP, 총대장균군 지표를 하위부문 지표로 두고 평균화된 값을 “유기물”이라는 새로운 대용 지표로 조정하여 종합지수 산정에 필요한 지표에 포함했다. 수생태 부문은 모든 조건을 충족시키는 것으로 나타났다.

최종적으로 총 4개 부문 15개 지표에 대해 지표를 재조정한 결과, 총 4개 부문 11개 지표가 생태적 건전성 지수를 위한 대표 지표로 선정되었다.

2. 하천 생태적 건전성 분석

1) 집계에 따른 가중치 기여율 분석

지표 선정 기준에 따라 수량, 수질, 수생태, 서식·수변 부문의 대표 지표를 산정하였으며 4개 부문을 다시 집계하여 생태적 건전성 지수를 산출하였다. 본 연구에서는 부문별 집계방식으로 기하 평균, 가중치는 동일 가중치를 적용하였다.

Figure 6은 수량, 수질, 수생태, 서식·수변 부문의 최종 집계에 따른 분석결과를 상관비와 그에 따른 기여율을 제시한 것이다. 부문별 지수가 생태적 건전성 지수에 미치는 기여율을 살펴보면, 수량 16.5%, 수질 26.0%, 수생태 40.6%, 서식·수변 16.9%로 수생태 지수가 가장 크게 기여한 것으로 나타났다. 즉, 동일

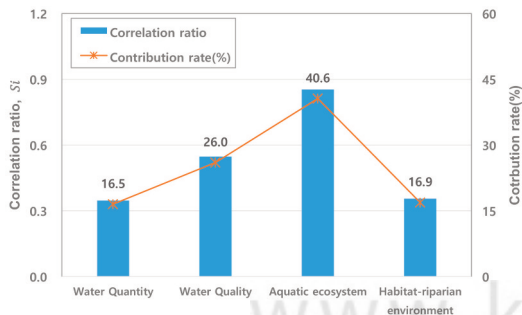


Figure 6. Results of S_i and Contribution rate.

가중치를 적용하더라도 집계(기하평균) 시 지수가 갖는 내재적 가중치에 의해 부문별 지수가 종합지수에 미치는 실제 영향은 다르게 나타나는 것으로 분석되었다. 따라서, 상관비 S_i 분석은 개별 지표가 종합지수에 실제 미치는 영향을 분석할 수 있어 지표의 구조적 특성뿐만 아니라 다양한 가중치의 적용에 따른 지수의 변화를 해석할 수 있는 유용한 방법인 것으로 판단된다.

2) 생태적 건전성 분석

산정된 생태적 건전성 지수의 결과 해석을 위하여 중권역 지점별 시간 변화에 따른 추세를 살펴보았다. 지표의 표준화 과정에서 시간에 따른 절대 변화 크기를 살펴보기 위하여 2008년을 기준년으로 지표를 표준화하였다. 2008년을 기준으로 지수 값이 100 이상일 때 건전성 지수가 ‘ 좋음’, 100 이하일 때는 상대적으로 ‘ 나쁨’으로 평가하였으며 기준년에 대한 시계열 변화의 절대 추세를 분석하여 중권역별로 하천의 생태적 건전성을 평가하였다.

기준년의 평균(건전성 지수 값: 100)에 따른 중권역별 지수 변화를 살펴보면, 용담댐(3001), 용담댐하류(3002), 무주남대천(3003), 영동천(3004), 초강(3005), 대청댐상류(3006), 보청천(3007), 대청댐(3008) 유역은 건전성 지수가 100 이상으로 ‘ 좋음’으로 미호천(3011), 금강공주(3012), 논산천(3013), 금강하구연(3014) 유역은 건전성 지수가 ‘ 나쁨’으로 분석되었다. 갑천(3009), 대청댐하류(3010)는 생태적 건전성이 좋아지다가 나빠지는 상태로 나타났다. 건전성 지수 값이 100 이상으로 좋은 상태를 유지하는

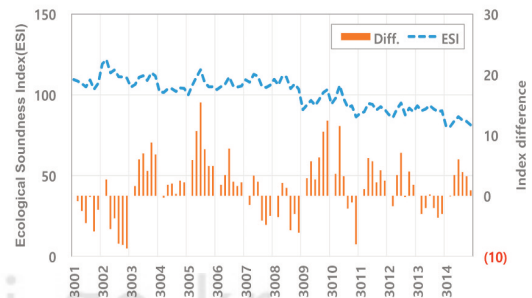


Figure 7. Results of ecological soundness index.

지역들은 댐 상류 유역(대청댐 기준)에 해당하는 중권역들이었으며 댐 하류 유역에 속하는 대부분 중권역은 지수 값이 100 이하로 건전성 지수가 나쁜 것으로 나타났다.

중권역별 시간 변화에 따른 건전성 지수 변화를 살펴보면, 용담댐, 용담댐하류, 보청천, 대청댐, 대청댐하류, 노산천은 시기별 다소간 차이는 있으나 건전성 지수가 악화하는 것으로 분석되었다. 반면 무주남대천, 영동천, 초강, 대청댐상류, 갑천, 미호천, 금강공주, 금강하구언은 시간 변화에 따라 개선되고 있는 것으로 분석되었다. 특히, 갑천 유역은 건전성 지수가 뚜렷하게 개선되고 있는 것으로 나타났으며 이는 하수처리장 및 비점오염원 개선사업에 지속적인 투자로 인한 결과로 판단된다.

VI. 결론 및 제언

지표의 투명성 및 일관성을 위해 탐색적 요인분석을 통해 1차 지표를 선정하여 지표 간 상관관계를 분석한 결과, 일부 지표는 여전히 상관관계가 높게 나타났다. 이러한 지표들은 집계(산술평균) 시 보상성 문제를 유발함으로써 새로운 지표로 재조정하였으며 지표 선정에서 음의 상관관계로 발생하는 음의 S_i 값은 나타나지 않았다. 각 지표가 지수에 갖는 영향 검토 결과, 대부분 큰 차이가 없었으며 지표별 특성을 잘 대변하는 것으로 나타났다.

선정된 지표들에 대해 동일 가중치를 적용하여 최종 집계(기하평균)한 생태적 건전성 지수에 대한 상관비 S_i 분석결과, 수생태 부문이 내재적 가중치에 의해 가장 큰 영향을 받는 것으로 나타났다.

생태적 건전성 지수의 분석결과, 댐상류 유역은 생태적 건전성이 좋은 것으로 나타났으며 댐하류 유역 대부분은 생태적 건전성이 나쁜 것으로 나타났다. 다만, 기준년(2008)에 대한 연도별 건전성 지수 변화를 살펴보면, 중권역별 특성에 따라 건전성 지수가 개선 또는 악화되는 경향을 나타내었다. 특히, 댐상류 유역 중 용담댐, 용담댐하류 유역은 건전성 지수가 전체적으로 좋음에도 불구하고 지수 값은 악화되는 것으로 나타났으며 댐하류 유역 중 갑천 유역은 건전성

지수가 크게 개선된 것으로 나타났다.

금강유역 14개 중권역에 대한 생태적 건전성은 유역의 특성에 따라 다양하게 나타나고 있으나 유역의 실제적 특성을 파악할 수 있는 기초자료는 많이 부족한 실정이다. 따라서 유역별 배출부하량, 오염 저감에 대한 투자비(하수처리장 신·증설, 비점오염 저감 시설 등), 택지개발 현황 등 유역별로 건전성 지수의 변화 원인을 파악할 수 있는 기초 정보 구축이 필요한 것으로 판단된다.

사사

이 논문은 2020학년도 충북대학교 연구년제 지원에 의하여 연구되었음.

References

- Becker W, Saisana M, Paruolo P, Vandecasteele I. 2017. Weights and importance in composite indicators: Closing the gap. *Ecological indicators* 80: 12-22.
- Becker W. 2017. Matlab toolbox for analysis and adjustment of weights in composite indicators, European Commission, Joint Research Centre, <http://dx.doi.org/10.13140/RG.2.2.27035.46883>.
- Diamantopoulos K, Riefler P, Roth KP. 2008. Advancing formative measurement models. *Journal of Business Research* 61: 1203-1218.
- Hauben M, Hung E, Hsieh WY. 2017. An exploratory factor analysis of the spontaneous reporting of severe cutaneous adverse reactions. *Therapeutic Advances in Drug Safety* 8(1): 4-16.
- Joo HG, Lee MG, Choi CG, Kim SJ, Kim HS. 2018. A Study on the selection of representative indicators of flood vulnerability assessment. *Journal of the Korean Society*

- of Hazard Mitigation 18(6): 335-346. [Korean Literature]
- Kim YS. 2016. Development of resilience indicator based on big data analysis under climate chang, Inha University, PhD thesis. [Korean Literature]
- Korea Water Resources Corporation. 2018. Study on the development a system for evaluating the health of the rivers comprehensively. KIWE-WSO-17-15. [Korean Literature]
- Lindén D. 2018. Exploration of implicit weights in composite indicators. KTH Royal institute of technology, Stockholm, Sweden.
- Mara TA, Tarantola S. 2012. Variance-based sensitivity indices for models with dependent inputs. Reliability Engineering & System Safety 107: 115-121.
- Mazziotta M, Pareto A. 2018. Measuring Well-Being Over Time: The adjusted Mazziotta-Pareto index versus other non-compensatory indices. Social Indicators Research 136: 967-976.
- Mazziotta M, Pareto A. 2016. On a generalized non-compensatory composite index for measuring socio-economic phenomena. Social Indicators Research 127: 983-1003.
- Munda G, Nardo M. 2005. Constructing consistent composite indicators: The Issue of Weights. EUR 21834 EN.
- OECD. 2008. Handbook on constructing composite indicators: methodology and user guide, available at <https://www.oecd.org/sdd/42495745.pdf>.
- Paruolo P, Saisana M, Saltelli A. 2013. Ratings and rankings: voodoo or science? Journal of the Royal Statistical Society: Series A, Statistics in Society 176L 609-634.
- Pinto U, Maheshwari B. 2011. River health assessment in Peri-urban landscapes: an application of multivariate analysis to identify the key variables. Water Res. 45: 3915-3924.
- Rowley HV, Peters GM, Lundie S, Moore SJ. 2012. Aggregating sustainability indicators: beyond the weighted sum. Journal of Environmental Management 111: 24-33.
- Saisana M, Saltelli A. 2011. Rankings and ratings: Instructions for use. Hague Journal on the Rule of Law 3: 247-268.
- Xu C, Gertner GZ. 2008. Uncertainty and sensitivity analysis for models with correlated parameters. Reliability Engineering & System Safety 93: 1563-1573.