

연구논문

## 회귀분석에 의한 TOC 농도 추정

- 오수천 유역을 대상으로 -

박진환 · 문명진 · 한성욱 · 이형진 · 정수정 · 황경섭 · 김갑순

국립환경과학원 영산강물환경연구소

## Application of Regression Analysis Model to TOC Concentration Estimation

- Osu Stream Watershed -

Jinhwan Park · Myungjin Moon · Sungwook Han · Hyungjin Lee ·  
Soojung Jung · Kyungsup Hwang · Kapsoon Kim

Yeongsan River Environment Research Center, National Institute of Environmental Research

### Abstract

The objective of this study is to evaluate and analyze Osu stream watershed water environment system. The data were collected from January 2009 to December 2011 including water temperature, pH, DO, EC, BOD, COD, TOC, SS, T-N, T-P and discharge. The data were used for principle component analysis and factor analysis. The results are as follows. The primary factors obtained from both the principal component analysis and the factor analysis were BOD, COD, TOC, SS and T-P. Once principal component analysis and factor analysis have been performed with the collected data and then the results will be applied to both simple regression model and multiple regression model. The regression model was developed into case 1 using concentrations of water quality parameters and case 2 using delivery loads. The value of the coefficient of determination on case 1 fell between 0.629 and 0.866; this was lower than case 2 value which fell between 0.946 and 0.998. Therefore, case 2 model would be a reliable choice. The coefficient of determination between the estimated figure using data which was developed to the regression model in 2012 and the actual measurement value was over 0.6, overall. It can be safely deduced that the correlation value between the two findings was high. The same model can be applied to get TOC concentrations in future.

Keywords : Factor Analysis, Osu Stream Watershed, Principal Component Analysis, Regression Analysis Model

## I. 서론

최근 생분해성 유기물을 나타내는 관리지표인 BOD(Biochemical Oxygen Demand)와 생분해성 유기물질을 포함한 난분해성 유기물질을 나타낼 수 있는 관리지표인 COD(Chemical Oxygen Demand) 및 TOC(Total Organic Carbon)의 상관성에 관한 많은 연구가 수행되고 있다(김범철 등, 2007; 박혜영 등, 2006; 변주태 등, 2008; 서희정 등, 2010). 또한 수질관리를 위해 국내의 경우 최근 들어 BOD와 COD가 반영하지 못하는 오염원의 총량이나 유기물의 특성이 반영된 TOC를 수질오염지표로 적용하기 위해 큰 관심을 기울이고 있다(박윤경 등, 2012). TOC의 경우 BOD와 COD에 비교해 볼 때 분석시간이 짧고, 분석오차를 줄일 수 있기에 관심이 커지고 있다(최동호 등, 2012). 한편, BOD와 COD를 이용하여 TOC 농도를 추정하기 위한 회귀모형의 개발과 평가에 대한 연구들이 있으며(민상운 등, 2012; 한상훈 등, 2012; 정재운 등, 2013), 유량-부하량 관계식을 이용하여 부하량 산정을 위한 연구가 많이 수행되고 있다(박지형 등, 2011; 박진환 등, 2012; 정상만 등, 2000; 최한규 등, 2002). 실제로 하천에 수질 특성을 파악하기 위해서는 많은 자료가 필요하고 이에 많은 예산과 인원이 투입되어야 할 것이며, 그럼에도 불구하고 측정이 불가능한 경우도 생길 것이다. 이처럼 결측 자료의 보완에 활용될 수 있는 농도 추정을 위한 연구는 앞으로도 지속되어야 할 것이며, 미래의 수질관리를 위한 중요한 분석기법이 될 것으로 판단된다. 그러나 앞서 언급해온 연구들은 회귀 분석을 수행하기 위해서 모집단의 변수들을 추출함에 있어서 대부분 Pearson 상관분석을 수행하였다. 이러한 방법은 두 변수 사이의 선형적 집중성만을 측정하게 되어 비선형의 연관성을 반영하지 못하며, 특이점에 따라서 분석결과가 크게 달라질 수 있다. 또한 상관계수를 이용하여 결과를 해석함에 있어 분석자의 판단에 의존해야 하기 때문에 분석결과에서 다소 차이가 발생할 수 있을 것이다. 한편 어떤 하천의 수질변화 및 특성을 해석하기 위한 연구로 다변량 분석기법인 주성분분석 및 요인분석이 최근 들어 많이

수행되고 있으며, 하천의 시·공간적 해석에 유용하게 쓰이고 있다(김영주, 2003; 김요용과 이시진, 2011; 박진환 등, 2013; 유순주 등, 2005; 유호식, 2002; 정승원 등, 2003). 이러한 방법 중에 주성분 분석은 요인분석을 수행하기 전에 몇 개의 인자들을 추출하기 위해서 여러 가지 방법이 사용되지만 그 중에 보편적으로 많이 사용하는 분석법이며, 요인분석은 수많은 변수들 중에 상관관계를 이용하여 여러 변수들로 측정된 자료를 보다 적은 수로 줄여 정보를 요약하고 변수들을 적절히 조합함으로써 묶여지지 않은 중요도가 낮은 변수를 제거할 수 있다. 이렇게 얻어진 요인들을 회귀분석에 이용한다면 보다 효율적으로 TOC 농도를 추정할 것으로 판단된다. 이에 본 연구에서는 주성분 및 요인분석의 결과로부터 추출된 수질 인자를 회귀모형의 독립변수로 하였으며, 종속변수는 TOC로 하여 TOC의 농도 추정을 위한 단회귀모형과 중회귀모형을 개발하였다. 또한 회귀모형은 Case 1(수질농도)과 Case 2(유달부하량)로 구분하여 개발한 후 적용성을 평가하였으며, 향후 오수천 유역의 수질관리정책 수립을 위한 기초자료를 제공하는데 목적이 있다.

## II. 연구방법

### 1. 연구대상 지역

섬진강 유역은 유역면적 4896.50 km<sup>2</sup>, 간선유로 연장 212.30 km이며, 국가하천 3개소로 하천 분류를 중심으로 하천연장은 173.30 km로 영산강과 탐진강에 비해 더 길다. 섬진강 유역의 주요하천은 국가하천 3개소(섬진강, 요천, 보성강)를 포함하며, 지방하천은 2개소(오수천, 추령천)이다. 본 연구의 대상 유역은 오수천 유역 중 발원지부터 섬진강분류점 전까지 전 구간 및 유역으로 유역면적은 371.1 km<sup>2</sup>이며, 측정 지점은 오수천 유역의 말단 지점이다(국립환경과학원, 2012).

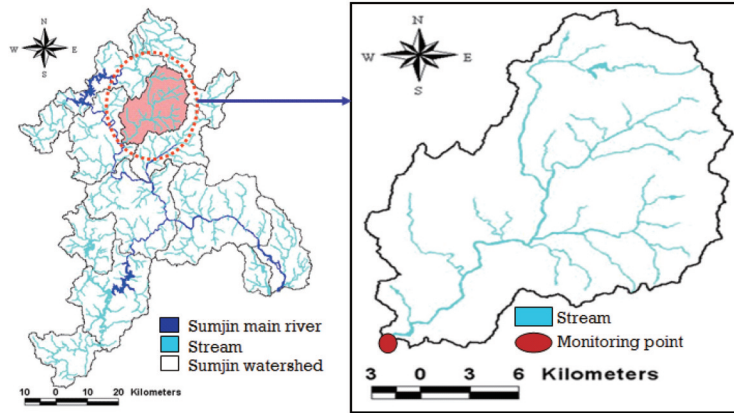


Figure 1. Location of study site and monitoring point

## 2. 자료조사 방법

수질 및 유량 조사는 4년(2009–2012)간 평균 8일 간격으로 실시하였으며, 대상유역 말단 지점에서 유량을 측정하고 동시에 수질시료를 채취하였다. 수질시료의 채취와 분석은 수질오염공정시험기준(환경부, 2008)에 준하여 실시하였으며, WT(Water Temperature), DO(Dissolved Oxygen), pH, EC(Electric Conductivity)는 YSI 650 MDS를 이용하여 현장에서 직접 측정하였다. 또한 하천유량은 수문 관측 매뉴얼(건설교통부, 2004)에 준하여 유속-면적법(Velocity-Area Method)으로 측정하였다.

## 3. 통계분석

3년(2009–2011)간 모니터링을 통해 측정된 자료로써 주성분분석(Principal Component Analysis)과 요인분석(Factor Analysis), 그리고 회귀분석(Regression Analysis)을 통계분석 프로그램인 SPSS(ver. 18.0)로 실시하였으며, 2012년의 측정 자료는 회귀모형의 적용성 평가에 활용하였다. 주성분 분석은 해석하고자 하는 변수 간의 상관성을 갖는 여러 종류의 데이터를 큰 손실 없이 요약하여 2 혹은 3 차원의 데이터로 축약하는 방법이며, 관측대상이 어떠한 위치에 있는지 시각적으로 파악할 수 있다. 요인분석은 다변량 분석법 중 하나로 여러 변수간의 상관행렬로부터 공통요인을 도출하여 서로 상관성이 있는 경우 변수간의 상관성을 설명하고 변수를 적은 수

로 축소시켜 요인을 추출하는 방법이다. 본 연구에서는 오수천 유역의 10개의 수질인자와 유량을 대상으로 주성분분석을 통하여 요인을 추출한 후 요인별 점수를 산정하는 요인분석을 실시하였다. 요인분석을 통해 분류된 제 1요인에 해당하는 인자들은 상관성이 높은 공통 인자들로서 회귀모형을 구축하기 위한 변수로 사용하였다. 회귀분석에는 단회귀분석(Simple Regression Analysis)과 중회귀분석(Multiple Regression Analysis)이 있으며, 설명변수가 하나인 경우가 단회귀분석이고 두 개 이상인 경우는 중회귀분석이다. 본 연구에서의 단회귀모형은 BOD와 COD를 독립변수로 식 (1)과 (2)의 방법으로 개발하였으며, 중회귀모형은 식 (3)과 같다.

$$y = b_0 + b_1x_1 \quad (1)$$

$$y = b_0 + b_2x_2 \quad (2)$$

여기서, 종속변수  $y$ 는 TOC 농도,  $b_1$ ,  $b_2$ 는 회귀계수,  $b_0$ 는 상수항,  $x_1$ ,  $x_2$ 는 독립변수로서 BOD와 COD의 농도이다.

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots + \beta_nx_n \quad (3)$$

여기서, 종속변수  $y$ 는 TOC 농도,  $x_i$ 는  $i$ 번째 독립변수,  $\beta_i$ 는  $y$ 에 대한  $i$ 번째 독립변수  $x_i$ 의 회귀계수이다.

회귀모형의 통계적인 유의성 검증은 F 검정법에 의한 F 통계량을 이용하여 계산된 F 통계량의 값에 대한 유의확률이 유의수준보다 작으면 회귀계수가

유의하다는 것으로 검정하였으며, 이와 같은 검정 방법은 한상훈 등(2012)과 정재운 등(2013)의 연구에서도 수행되었다.

개발된 회귀모형을 이용하여 2012년의 측정 자료를 대상으로 TOC를 추정하였으며, 추정된 값과 실측값의 평가는 대응표본 T검정을 통해 검정하였다. 여기서 T검정에 대한 유의확률(양쪽)이  $p < 0.05$ 이면 추정 값과 실측값 사이에 유의성이 있는 것으로 볼 수 있으며, 이는 두 값이 동일하지 않고 차이가 있는 것으로 해석하며 모형의 신뢰도가 낮다는 것을 의미한다. 한편 유의확률(양쪽)이  $p > 0.05$ 이면 추정 값과 실측값 사이에 유의성이 없는 것으로 보며, 회귀모형의 신뢰도가 높은 것으로 해석할 수 있다.

### III. 결과 및 고찰

#### 1. 조사지점의 수질 특성

연구지점으로부터 3년(2009-2011)간의 자료와 2012년에 실측된 자료의 분석결과는 Figure 2 및 Table 1과 2에 나타내었다. 3년 동안 측정된 BOD, COD, TOC의 농도 범위는 각각 0.5-6.1 mg/L에서 평균 1.6 mg/L이며, 2.0-11.2 mg/L, 4.5 mg/L와 1.3-10.2 mg/L, 3.2 mg/L로 나타났다. T-N과 T-P의 경우 0.6-3.9 mg/L의 농도 범위에서 평균 1.8 mg/L와 0.012-0.665 mg/L의 농도범위에서 평균 0.058 mg/L로 나타났으며, SS는 농도 범위는 0.12

-38.2 mg/L이며, 평균 11.5 mg/L로 조사되었다. 2012년의 경우 BOD, COD, TOC의 농도 범위는 0.6-6.1 mg/L이고 평균 1.8 mg/L와 2.7-14.3 mg/L, 평균 5.0 mg/L, 0.7-6.5 mg/L, 평균 2.9 mg/L로서 3년 평균 농도와 비교해 볼 때 BOD와 COD는 높았으며, TOC는 낮게 조사되었지만 큰 차이는 없는 것으로 나타났다. 또한 T-N과 SS는 농도 범위는 각각 0.5-2.8 mg/L, 평균 1.9 mg/L와 1.0-278.8 mg/L, 13.4 mg/L로서 다소 높게 조사되었으며, T-P의 경우 0.042-0.273 mg/L의 농도 범위에서 평균 0.048 mg/L로 다소 낮게 조사되었다. 실측된 3년(2009-2011)간의 자료와 2012년의 자료는 다소 차이는 있지만 큰 차이는 보이지 않았으며, 본 연구지점에 대한 BOD의 하천 생활환경기준은 Ib등급(좋음)이며, COD와 T-P의 경우 II등급(약간 좋음)에 해당한다.

#### 2. 주성분 및 요인분석

회귀모형의 독립변수로 사용하기 위하여 대상지점에 대한 주성분분석을 실시하였고 추출된 요인으로부터 상관성이 높은 변수들로 분류하여 Table 3, 4에 나타내었다. 주성분 수를 결정하기 위해 요인을 설명하는 분산의 크기를 나타내는 고유치(Eigenvalue)를 기준으로 통상 고유치가 1이상인 값을 기준으로 결정하였다. 고유치는 요인을 설명하는 분산의 크기를 나타내며 고유치가 1보다 크다는 것은 하나의 요인이 변수 1개 이상을 설명해 주며, 1보다

Table 1. Descriptive statistics of water quality parameters (2009-2011)

Variable	Discharge	WT	pH	EC	DO	BOD	COD	SS	T-N	T-P	TOC
Maximum	114.469	29.0	8.8	433.0	14.1	6.8	11.2	238.2	3.9	0.665	10.2
Minimum	0.217	0.0	6.8	90.0	6.7	0.5	2.0	0.1	0.6	0.012	1.3
Median	3.282	17.5	7.9	200.0	11.5	1.4	3.9	5.0	1.8	0.038	2.8
Average	7.566	16.2	7.8	198.0	11.3	1.6	4.5	11.5	1.8	0.058	3.2

Table 2. Descriptive statistics of water quality parameters (2012)

Variable	Discharge	WT	pH	EC	DO	BOD	COD	SS	T-N	T-P	TOC
Maximum	105.457	32.9	8.7	234.0	14.8	6.1	14.3	278.8	2.8	0.273	6.5
Minimum	0.477	1.7	6.6	98.0	7.5	0.6	2.7	1.0	0.5	0.011	0.7
Median	4.226	18.3	7.8	180.5	11.0	1.3	4.1	5.7	1.9	0.042	2.6
Average	10.296	16.7	7.8	168.8	11.3	1.7	5.0	13.4	1.9	0.048	2.9

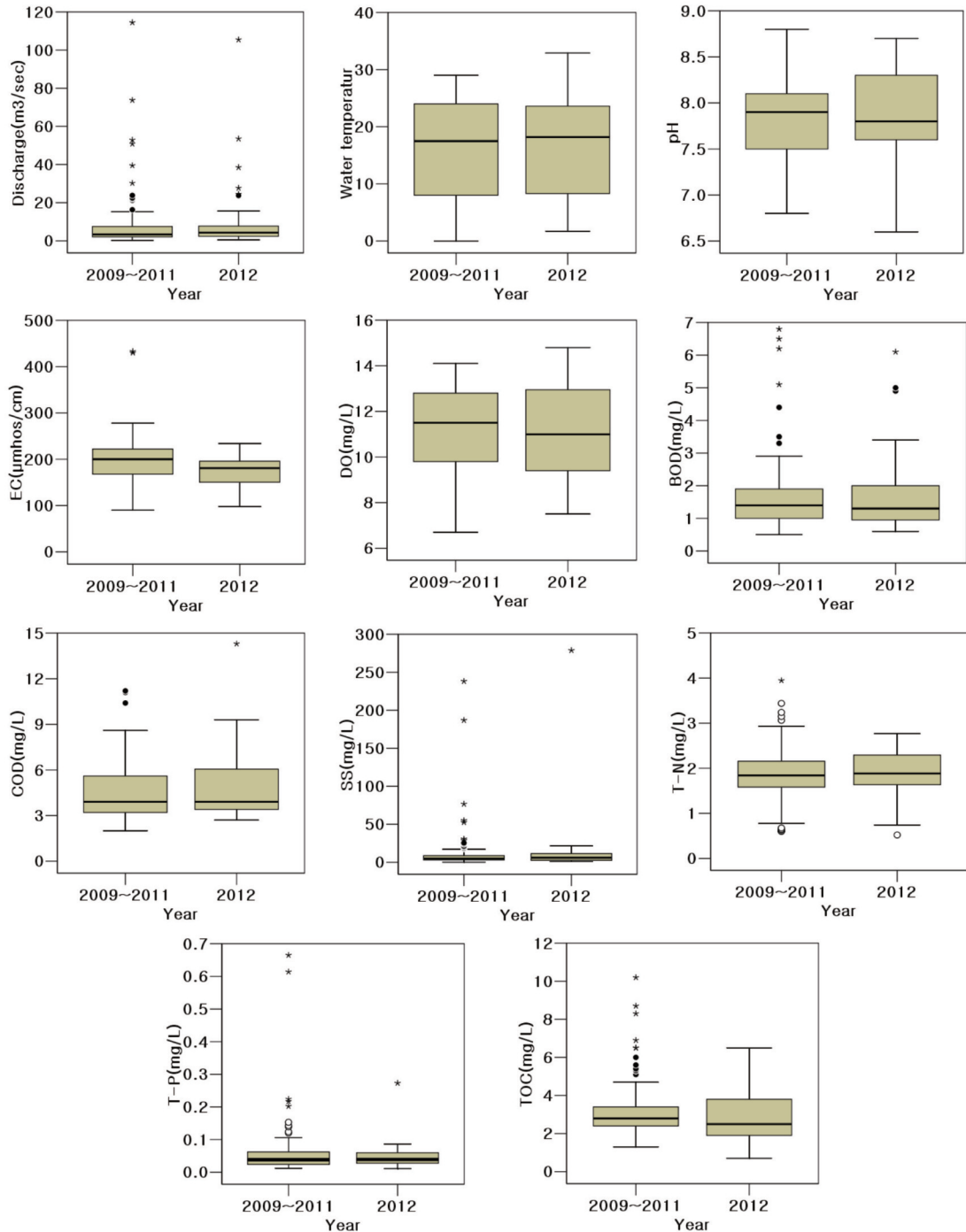


Figure 2. Box plot of major water quality parameters in Osu stream watershed

작다는 것은 변수 1개의 분산도를 설명할 수 없다는 것으로 해석할 수 있다. 또한 변수들의 요인 부하량이 어느 한 요인에 높게 나타나도록 하기 위한 방법

인 Varimax 회전방법을 적용하였다. 연구 지점에서 3개의 요인이 추출되었으며, 제 1요인의 고유치와 기여율은 5.504, 50.037%, 제 2요인과 3요인은 각각

Table 3. Eigen values and cumulative percentage of factors

Component	Initial eigen values			Extraction sums of squared loadings		
	Total	Variance(%)	Cumulative(%)	Total	Variance(%)	Cumulative(%)
1	5.504	50.037	50.037	5.504	50.037	50.037
2	1.628	14.800	64.837	1.628	14.800	64.837
3	1.186	10.785	75.622	1.186	10.785	75.622
4	0.976	8.869	84.491			
5	0.793	7.209	91.700			
6	0.428	3.887	95.587			
7	0.200	1.817	97.404			
8	0.106	0.963	98.367			
9	0.094	0.856	99.223			
10	0.065	0.593	99.815			
11	0.020	0.185	100.000			

Table 4. Rotated component matrix factor analysis

Parameters	Component		
	1	2	3
Discharge	0.605	<b>0.608</b>	0.177
WT	0.226	0.032	<b>0.934</b>
pH	-0.196	<b>-0.686</b>	0.063
EC	0.098	<b>-0.615</b>	-0.234
DO	-0.379	-0.342	<b>-0.750</b>
BOD	<b>0.891</b>	0.080	0.051
COD	<b>0.874</b>	0.001	0.189
TOC	<b>0.924</b>	0.024	0.225
SS	<b>0.796</b>	0.433	0.030
T-N	0.415	0.492	-0.487
T-P	<b>0.804</b>	0.447	0.146

1.628, 14,800%와 1.186, 10,785%이고 누적기여율은 75.622%로서 전체 수질변동에 대하여 설명하고 있다. 한편 전체 기여율의 50%이상을 설명해주는 제 1요인은 BOD, COD, TOC, SS, T-P로 분류되었으며, 제 2요인과 3요인은 각각 유량, pH, EC와 WT, DO로 분류하였다. 이중 제 1요인의 항목을 대상으로 회귀모형에 적용하였다.

### 3. 회귀모형의 개발 및 검증

단회귀모형에서 실측된 BOD와 COD농도는 각각의 독립변수를 가진다. 또한, 중회귀모형의 경우 상수항과 BOD, COD, SS, T-P와 같은 독립변수를 가

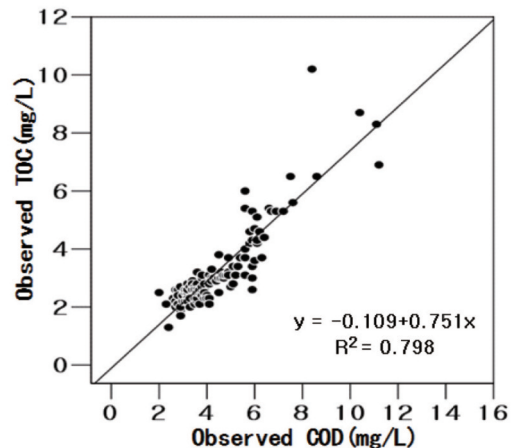
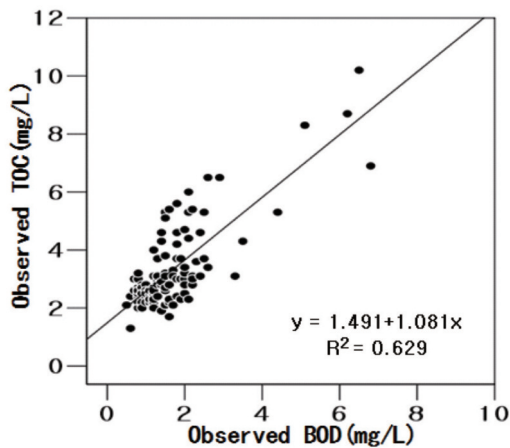


Figure 3. Simple regression model of case 1

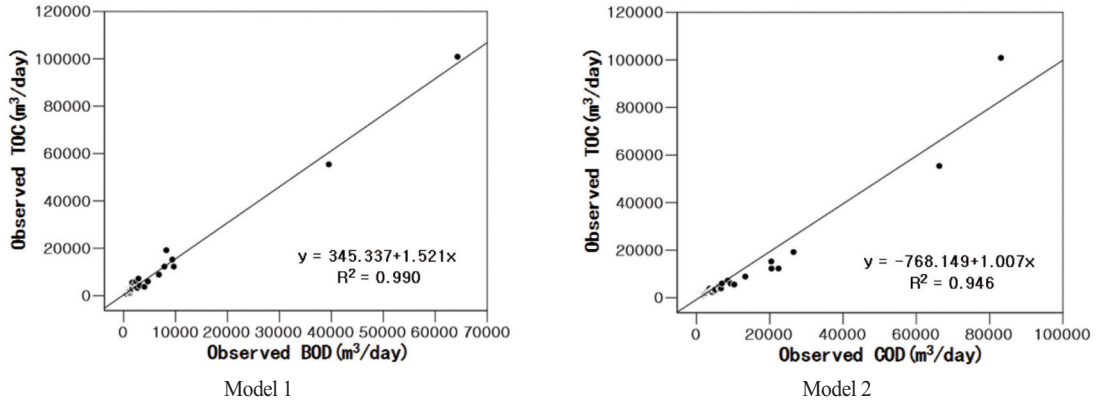


Figure 4. Simple regression model of case 2

Table 5. Selected the simple and multiple regression model for TOC estimation

Type	Model	Regression models	r	R <sup>2</sup>	F-value	p-value
Case 1	1	TOC = 1.491 + 1.081*BOD	0.793	0.629	208.9	0.000
	2	TOC = -0.109 + 0.751*COD	0.894	0.798	487.3	0.000
	3	TOC = 0.359 + 0.006*BOD + 0.587*COD + 0.009*SS + 2.585*T-P	0.930	0.866	193.3	0.000
Case 2	1	TOC = 345.337 + 1.521*BOD	0.995	0.990	11897.4	0.000
	2	TOC = -768.149 + 1.007*COD	0.973	0.946	2151.9	0.000
	3	TOC = 258.124 + (-0.100*BOD) + 0.607*COD + 0.043*SS + (-7.290*T-P)	0.999	0.998	14484.7	0.000

지게 되며, TOC의 농도는 종속변수이다. 본 연구에서는 수질농도만을 대상으로 개발된 Case 1과 수질농도에 유량을 곱하여 산정된 유달부하량(Delivery Load)으로 개발된 Case 2로 구분하였으며, 그 결과는 Figure 3, 4와 Table 5에 나타내었다. 개발된 모든 모형들의 F-value에 대한 p-value가 0.000으로 나타났으며, 이는 개발된 모형의 종속변수인 TOC를 설명하는데 유의한 것으로 볼 수 있다. 한편, Case 1의 Model 3의 결정계수(Coefficient of Determination, R<sup>2</sup>)가 0.866으로 가장 높았으며, Case 2의 경우 모든 모형의 R<sup>2</sup>가 0.9이상으로 나타나 Case 1의 모형보다 신뢰성이 높은 것으로 사료된다.

#### 4. 회귀모형의 적용성 평가

회귀모형은 3년(2009-2011)간 실측된 자료를 이용하여 개발되었다. 개발된 단회귀모형과 중회귀모형의 적합성 평가를 위하여 2012년의 실측자료를 활용하여 TOC농도를 추정하였으며, 실측치와 추정치

에 대한 분석 결과는 Figure 5와 Table 6에 나타내었다. 회귀모형에 의해 산정된 추정치와 실측치의 결정계수는 Case 1의 Model 1은 0.594로서 가장 낮았으며, Model 1과 2는 0.8이상으로 나타냈다. Case 2(Model 1, 2, 3)의 경우 0.9이상으로 측정치와 추정치의 상관성이 매우 높은 것으로 볼 수 있으며, 유달부하량을 이용한 Case 2의 회귀모형이 TOC농도를 추정함에 있어 수질농도를 이용한 Case 1의 회귀모형보다 높은 상관성을 나타냈다. 한편, 대응표본 T검정을 통해 TOC의 측정치와 추정치 간의 차이를 신뢰구간 95%에서 검정을 하였으며, 그 결과 Case 1(Model 1, 2, 3)의 경우 양쪽 유의확률이 p<0.05로 나타나 두 값의 차이는 유의하다고 판단하였으며, 반면 Case 2(Model 1, 2, 3)의 경우 양쪽 유의확률에서 p>0.05로서 두 값의 차이가 유의하지 않은 것으로 볼 수 있다. 즉, 통계적으로 유의하다는 것은 두 값의 차이가 있는 것으로 볼 수 있으며, 유의하지 않다는 것은 두 값의 차이가 없다는 것으로 해석할 수 있다.

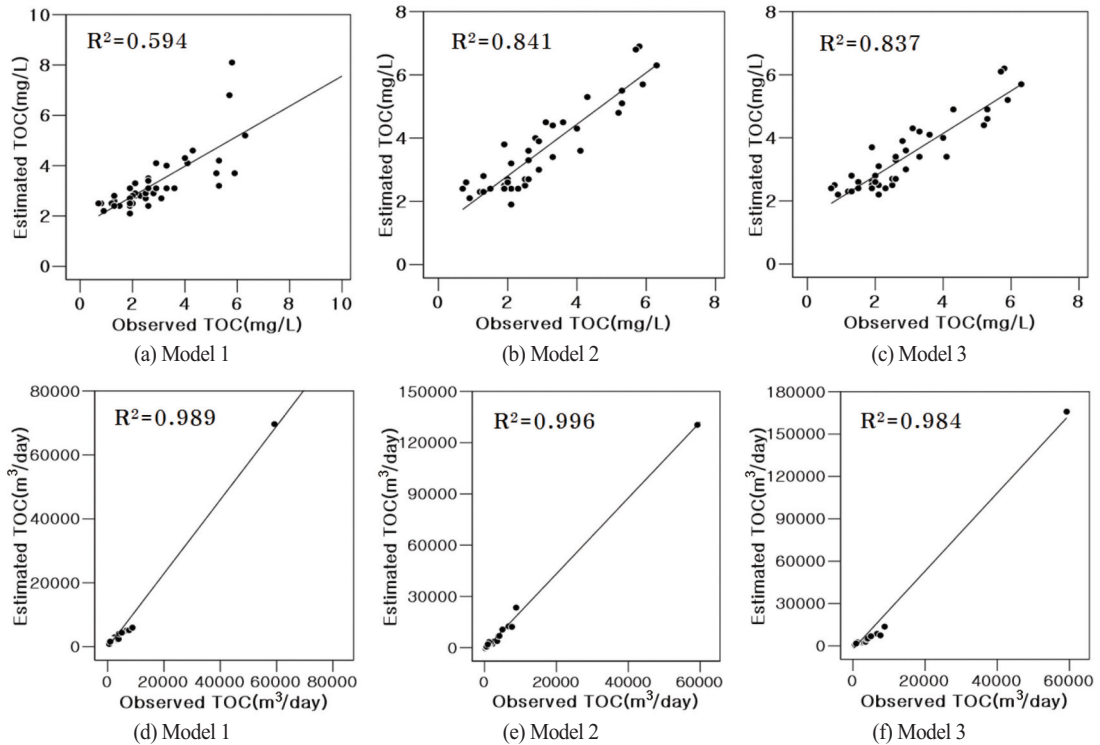


Figure 5. Comparison of the observed and estimated TOC using regression models (Case 1: a, b, c and Case 2: d, e, f)

Table 6. Estimation of the observed and estimated TOC using paired T-test

Model	Case 1	Case 2
	p-value	p-value
1	0.003(p<0.05)	0.344(p>0.05)
2	0.000(p<0.05)	0.150(p>0.05)
3	0.000(p<0.05)	0.239(p>0.05)

실측치와 추정치에 대한 모형의 적합성 평가에서 단순히 결정계수로 모형을 평가한다면 Case 1과 Case 2의 회귀모형은 TOC 농도를 추정함에 있어 적용이 가능할 것으로 판단되었다. 하지만, 대응표본 T검정을 이용한 검정 결과에서 Case 1 보다는 Case 2의 회귀모형을 이용하여 TOC 농도(유달부하량)를 추정하는 것이 더욱 신뢰성이 높을 것으로 판단하였다.

#### IV. 결론

본 연구는 오수천 유역에서 실측된 자료를 활용하여 주성분 및 요인분석을 실시하였으며, 추출된 인자

들을 통해 회귀모형을 개발하고 적용성 평가를 수행하여 다음과 같은 결론을 도출하였다. 주성분 및 요인분석에서 전체 기여율의 50%이상을 설명해주는 제 1요인은 BOD, COD, TOC, SS, T-P로 분류되었으며, 제 1요인의 수질농도만을 대상으로 개발된 Case 1(Model 1, 2, 3)에 대한 회귀모형의 결정계수는 0.629-0.866으로 유달부하량을 대상으로 개발된 Case 2(Model 1, 2, 3)의 0.946-0.998보다 낮게 나타나 Case 2의 모형이 신뢰성이 더 높은 것으로 판단할 수 있다. 또한 개발된 모든 모형의 F-value에 대한 p-value가 0.000으로 나타나 종속변수 TOC를 설명하는데 유의한 것으로 판단하였다. 개발된 회귀모형을 2012년 자료에 적용하여 산정된 추정치와 실측치에 대한 결정계수는 Case 1의 Model 1을 제외하고 전반적으로 0.8이상이며, 두 값의 상관성이 높은 것으로 볼 수 있다. 한편, 대응표본 T검정을 통해 TOC의 측정치와 추정치 간의 차이를 신뢰구간 95%에서 해석해보면 Case 1(Model 1, 2, 3)의 경우 양쪽

유의확률이  $p < 0.05$ 로 나타나 두 값의 차이는 유의하다고 판단하였으며, 반면 Case 2(Model 1, 2, 3)의 경우 양쪽 유의확률에서  $p > 0.05$ 로서 두 값의 차이가 유의하지 않는 것으로 해석하였다.

향후, Case 1(Model 2, 3)과 Case 2(Model 1, 2, 3)의 회귀모형은 TOC농도를 추정함에 있어 적용이 가능할 것으로 판단되며, 그중 유달부하량을 이용하여 개발된 Case 2의 회귀모형을 적용한다면 좀 더 신뢰도 높은 값을 추정할 수 있을 것으로 판단된다. 그러나 본 연구결과는 오수천 유역에서만 나타나는 특성에 주의할 필요가 있으며, 다른 유역에 적용하기 위해서는 TOC와 상관성이 없는 인자들과 더불어 실험분석에 영향을 미치는 인자들의 관계에 대해서도 추가적인 연구가 필요할 것으로 사료된다.

## 참고문헌

- 건설교통부, 2004, 수문관측매뉴얼.
- 국립환경과학원, 2012, 영산강·섬진강수계 유량 측정망 운영사업 최종보고서.
- 김병철, 정성민, 장창원, 김재구, 2007, 호수와 하천에서 유기물 오염도의 지표로서 BOD, COD와 TOC의 비교 및 분해율 산정, 대한환경공학회지, 29(6), 640-643.
- 김영주, 2003, 주성분분석에 의한 도시호수의 수질 평가에 관한 연구, 환경관리학회지, 9(2), 197-203.
- 김요용, 이시진, 2011, 다변량 통계 분석기법을 이용한 한강수계 지천의 수질 평가, 대한환경공학회지, 33(7), 501-510.
- 민상윤, 이승필, 김진식, 박종운, 김만수, 2012, 하수처리장 방류수 수질예측을 위한 다중회귀 분석 모델 개발 및 검증, 대한환경공학회지, 34(5), 312-315.
- 박윤경, 이재관, 박무중, 김상단, 2012, 낙동강 최상류유역의 토지 이용별 TOC 부하량 산정, 한국물환경학회·대한상하수도학회 공동학술발표회 논문집, 736-737.
- 박지형, 손수민, 류덕희, 2011, 경험식과 요인분석을 통한 오염물질 유출 특성 연구 -남한강 상·하류 수계 주요 하천을 중심으로-, 수질보전 한국물환경학회지, 27(6), 905-913.
- 박진환, 정재운, 김대영, 김갑순, 한성욱, 김현욱, 임병진, 2013, 통계분석을 이용한 영산강·섬진강수계 주요 유입지천의 수질 특성, 한국환경영향평가학회지, 22(2), 171-181.
- 박진환, 임병진, 정재운, 김대영, 오태운, 이동진, 김갑순, 2012, 유량변동에 따른 탐진 A와 B 유역에서의 오염물질 유출 특성, 한국환경영향평가학회지, 21(6), 917-925.
- 박혜영, 이중기, 하훈, 이호변, 김양기, 박찬오, 박송인, 2006, 수질관리기술 ; TOC를 이용한 유기물질의 상관성 연구, 대한환경공학회 춘계학술발표회 논문집, 834-842.
- 변주태, 김태동, 이석기, 홍태기, 김현욱, 2008, 환경수 중 유기물량 지표로서 TOC, UVA 및 COD(Mn)간의 상관관계, 한국환경분석학회지, 11(4), 286-291.
- 서희정, 강영주, 민경우, 이경석, 서광엽, 김승호, 백계진, 김성준, 2010, 하천수와 하수처리장 방류수의 유기물 분포 및 분해 특성, 한국분석과학회지, 23(1), 36-44.
- 유순주, 채민희, 황종연, 이재안, 박종겸, 최태봉, 2005, 담수 이후 용담호 영양상태 변동 요인 분석, 수질보전 한국물환경학회지, 21(4), 360-367.
- 유호식, 2002, 한강수질에 영향을 끼치는 요인들의 통계분석, 대한환경공학회지, 24(12), 2139-2150.
- 정상만, 임경호, 최정현, 2000, 금강지류 유역에서의 유출량과 오염부하량의 상관관계 분석, 한국수자원학회논문집, 33(5), 527-536.
- 정승원, 이진환, 유종수, 2003, 한강하류의 환경학적 연구 v. 식물플랑크톤 군집의 대발생 특징, 한국조류학회지, 18(4), 255-262.
- 정재운, 조소현, 최진희, 김갑순, 정수정, 임병진, 2013, 하천수내 TOC 농도 추정을 위한 단순 회귀모형과 다중회귀모형의 개발과 평가,

- 한국물환경학회지, 29(5), 625-629.
- 최동호, 정재운, 윤광식, 이경숙, 최우정, 임상선, 박하나, 임병진, 황태희, 2012, 영농기간동안 논 유출수 BOD, COD, 자료를 이용한 TOC 농도 추정, 한국물환경학회지, 28(6), 813-818.
- 최한규, 백경원, 최용물, 오기호, 2002, 유량변동에 따른 소양강무역 수질의 통계학적 해석, 대한토목학회논문집, 22(2), 117-124.
- 한상훈, 이은형, 김현욱, 2012, 소프트웨어 센서를 이용한 수체내 총질소 총인 농도 예측, 한국환경분석학회지, 15(3), 215-223.
- 환경부, 2008, 수질오염공정시험법.