

한강하류부 수질자료의 추이분석에 관한 연구

Trend Analysis for Water Quality Time Series in the Downstream of Han River

○백중철*, 허준행**, 조원철***

1. 서론

수질 시계열 자료는 비정규분포, 계절성(seasonality), 유량과의 상관성, 결측자료, 검출한계이하의 자료, 계열상관 등의 특성을 포함하므로 분석이 난해하다. 본 연구에서는 비정규분포, 계절성, 결측자료 그리고 검출한계 이하의 자료를 포함하는 수질자료의 추이 분석에 적합한 비매개변수 검정기법(nonparametric test), 즉 계절적 Kendall 검정기법과 추이의 크기를 추정하기 위한 계절적 Kendall 변화를 추정기법을, 실측 수질자료에 적용하여 수질시계열자료의 추이 및 관측지점간 수질자료 추이의 동질성 등에 대한 통계적 특성을 분석하였다.

한강수계내에서의 수질관측 자료중 오염도가 비교적 심한 팔당댐 이하 하류부에 위치한 11개 수질 관측지점의 자료를 대상으로 관측지점 사이의 수질자료의 동질성 분석 등의 기본적인 통계학적 특성 및 수질자료의 추이를 분석하였다.

2 기본이론

2.1 계절적 Kendall 검정기법(Seasonal Kendall Test)

Hirsch 등(1982)은 계절성(seasonality)의 존재에 대해서 민감하지 않은 계절적 Kendall 검정기법을 제안하였다. 이 검정법에서 귀무가설(null hypothesis) H_0 '은 대립가설(alternative hypothesis) H_0 '의 완화된 형태이다. X 를 하부표본 X_1 부터 X_{12} 까지로 구성된 전체표본이라하고 하부표본 X_i 는 i 월로부터 n_i 개의 연간 자료를 포함한다고 하면, 계절성 Kendall 검정법에 대한 귀무가설 H_0 '은 X 가 독립확률변수 $(x_{i,j})$ 의 표본이며 X_i 가 독립적이고 동일하게 분포된 확률변수 $i = 1, \dots, 12$ 의 하부표본이라는 것이다. 추정치 S_i 는 다음과 같이 정의된다.

$$S_i = \sum_{k=1}^{n_i-1} \sum_{j=k+1}^{n_i} \text{sgn}(x_{ij} - x_{ik}) \quad (1)$$

H_0 '하에서 하부표본 X_i 는 Mann 검정법에서의 귀무가설 H_0 를 만족시킨다. S_i 에 대한 평균과 분산은 다음과 같이 정의되며, S_i 는 $n_i \rightarrow \infty$ 에 대해서 정규분포이다.

$$E[S_i] = 0 \quad (2)$$

$$\text{Var}[S_i] = \frac{n_i(n_i - 1)(2n_i + 5) - \sum_{t=1}^{n_i} t_i(t_i - 1)(2t_i + 5)}{18} \quad (3)$$

한편 $S' = \sum_{i=1}^{12} S_i$ 에 대한 기대치 및 분산은 다음과 같다.

* 연세대학교 공과대학 토목공학과 박사과정

** 연세대학교 공과대학 토목공학과 부교수

*** 연세대학교 공과대학 토목공학과 교수

$$E[S'] = \sum_{i=1}^{12} E[S_i] = 0 \quad (4)$$

$$\text{Var}[S'] = \sum_{i=1}^{12} \text{var}[S_i] + \sum_{i=1}^{12} \sum_{j=1, j \neq i}^{12} \text{cov}(S_i, S_j), \quad i \neq j \quad (5)$$

여기서 S_i 와 S_j ($i \neq j$)은 독립확률변수의 함수이므로 $\text{cov}(S_i, S_j) = 0$ 이다. 따라서 식 (5)의 $\text{Var}[S']$ 는 식 (6)과 같이 표현되며, 정규분포화에 의한 표준정규변량 Z' 은 식 (7)과 같이 정의된다.

$$\text{Var}[S'] = \sum_{i=1}^{12} \text{var}[S_i] = \sum_{i=1}^{12} \frac{n_i(n_i - 1)(2n_i - 5)}{18} \quad (6)$$

$$Z' = \begin{cases} \frac{S' - 1}{(\text{Var}(S'))^{1/2}} & \text{if } S' > 0 \\ 0 & \text{if } S' = 0 \\ \frac{S' + 1}{(\text{Var}(S'))^{1/2}} & \text{if } S' < 0 \end{cases} \quad (7)$$

추이의 존재에 대한 대립가설 H_1 대한 무(無) 추이의 귀무가설 H_0 를 검정하는데 있어서, 양측기각역검정(two-tailed test)의 경우, 만약 $|Z| \leq z_{\alpha/2}$ 이면 H_0 는 수용된다. 여기서 $F_N(z_{\alpha/2}) = \alpha/2$ 이며, F_N 은 표준정규누가분포함수이고, α 는 검정에 대한 유의수준(significance level)이다. 양의 S' 값은 시간에 따라서 값이 증가하는 증가추이(upward trend)를 그리고 음의 S' 값은 감소추이(downward trend)를 의미한다.

2.2 계절적 Kendall 변화율 추정량(Seasonal Kendall Slope Estimator)

계절적 Kendall 변화율 추정법은 계절성을 고려하기 위해서 Thell(1950)과 Sen(1968)이 제안한 방법을 일반화 한 것이다. 추이를 나타내는 시계열 자료를 분석하기 위해서는 추이의 크기(magnitude)를 추정하여야 한다. 이 크기는 변화율(시간당 변화량)로서 나타내어진다.

계절적 Kendall 변화율 추정량 B 는 먼저 모든 i 개의 계절(season)에 대해서 다음의 식(8)과 같이 d_{ijk} 값을 계산한다. 여기서 x_{ij} , x_{ik} ($i = 1, 2, \dots, 12$; $1 \leq k < j \leq n_i$)에 대해서 계산된 d_{ijk} 값들 중 중간값(median)이 계절적 Kendall 변화율 추정량이다.

$$d_{ijk} = \frac{x_{ij} - x_{ik}}{j - k} \quad (8)$$

추정량 B 는 계절적 Kendall 검정 통계량 S' 과 관계가 있다. 즉, S' 는 양(+)의 d_{ijk} 값들의 개수에 서 음(-)의 d_{ijk} 값들의 개수를 뺀 값이고 B 는 d_{ijk} 값중 중간값이므로, 만약 $S' > 0$ 이면 $B \geq 0$ 그리고 $S' < 0$ 이면 $B \leq 0$ 이 된다.

2.3 동질성 분석

수질자료가 동일한 유역내의 둘 이상의 관측지점에서 수집된 경우, 유역 전반에 대한 추이를 분석하는 것은 중요하다. 만약 동일 유역내의 관측지점에서의 추이가 동일한 경우 즉, 모두가 상승 또는 하강추이인 경우, 단순추이의 존재유무를 파악하는 것은 의미가 있으며, 이러한 분석은 서로 다른 관측지점에서 관측된 수질자료의 추이의 동질성을 분석함으로써 이루어진다. 추이의 동질성분석은 χ^2 통계량으로 실시하게 된다. 먼저 m 관측지점의 i 월의 Mann-Kendall 통계치 S_{im} 을 계산한 후 다음과 같이 정의되는 Z_{im} 을 계산한다.

$$Z_{im} = \frac{S_{im}}{[\text{VAR}(S_{im})]^{1/2}}, \quad i=1, 2, \dots, K, \quad m=1, 2, \dots, M \quad (9)$$

여기서 $\text{VAR}(S_{im})$ 은 식(6)으로 정의되며, μ_{im} 은 식(4)로 정의되는 통계치를 계산한다.

$$\bar{Z}_{i.} = \frac{1}{M} \sum_{m=1}^M Z_{im}, \quad i=1, 2, \dots, K \quad (10.1)$$

$$\bar{Z}_{.m} = \frac{1}{K} \sum_{i=1}^K Z_{im}, \quad m=1, 2, \dots, M \quad (10.2)$$

$$\bar{Z}_{..} = \frac{1}{KM} \sum_{i=1}^K \sum_{m=1}^M Z_{im} \quad (10.3)$$

다음은 식(11)로 정의된 $\chi^2_{station}$ 을 계산함으로써 관측지점간의 수질자료의 동질성을 분석할 수 있다.

$$\chi^2_{station} = K \sum_{m=1}^M \bar{Z}_{.m}^2 - KM \bar{Z}_{..}^2 \quad (11)$$

$\chi^2_{station}$ 는 chi-square 표에서 $M-1$ 의 자유도를 가지는 유의수준 α 의 한계값이다. 만약 $\chi^2_{station}$ 이 무의미(nonsignificant)하다면, 각 관측지점별로 서로 다른 추이를 분석하여야 한다.

3. 기법의 적용

본 연구에서는 한강 수계 수질관측소중 팔당댐으로부터 한강의 종점인 김포사이에 위치하고 있는 11개 수질관측소에서 관측된 DO, BOD, COD, SS, pH에 대한 9년간의 월별 수질자료에 대해서 추이분석 및 수질자료의 통계학적 특성을 분석하였다.

5개 항목의 수질자료에 대해서 Chi-Square 검정을 실시하여 동질성을 분석한 결과와 계절적 Kendall 검정기법과 계절적 Kendall 변화율 추정기법을 이용하여 관측지점별로 수질자료의 추이를 분석한 결과를 요약하면 표 1와 같다.

관측지점사이의 동질성 분석을 위한 Chi-Square 검정결과를 보면, 관측소간의 수질자료에 대한 Chi-Square 검정통계값의 경우, 자유도가 1이고 유의수준이 0.05이면 Chi-Square 분포에서의 한계값은 3.84으로서, 인접한 관측지점사이의 수질자료가 이 한계값을 초과하는 확률을 산정한 결과는 표 1에서 $P(\chi^2)$ 의 값에 해당한다. 이 한계값을 초과하는 경우는 즉, BOD를 예로 들어보면, 팔당-구리, 보광-노량진, 영등포-가양, 가양-행주의 수질관측자료에 해당한다. 이러한 경우 대상이 되는 두 개의 관측소에서의 추이는 동일한 방향이 아닌 서로 다른 추이의 가지고 있음을 의미한다.

한편, 계절적 Kendall 검정기법을 적용한 결과를 보면, $|Z| \leq z_{\alpha/2}$ 의 확률이 0.0 또는 0.0에 가까운 값은 귀무가설을 기각하고 대립가설을 수용하는 즉, 추이가 존재하는 것을 나타낸다. 표 1에서와 같이 DO의 경우 팔당 1, 노량진, 가양지점, BOD의 경우 구리, 영등포, 노량진, 가양, 행주, 김포지점, COD의 경우는 팔당 2, 영등포, 가양, 김포, SS는 구의 이하에 위치한 수질관측소 모두, 그리고 pH는 팔당 1, 구리, 구의, 영등포, 가양의 경우가 뚜렷한 추이를 보이는 것으로 나타났다. 추이의 크기 즉, 변화율(Unit/Years)은 표 1에서 제시한 바와같이 나타났으며, LL과 UL에 해당하는 값은 각각 유의수준 0.05에 해당하는 하한계 및 상한계 값이다.

한편, 관측된 수질자료의 월별 분포특성을 파악하기 위해서 BOD의 경우를 예로 하여 수질자료를

월별로 도시하고 평균 및 표준편차를 나타내면 그림 1과 같다. 그림 1을 분석해 보면 구리, 구의, 뚝도, 보광, 영등포, 노량진의 경우는 뚜렷한 계절적 변동특성을 보이는 것으로 나타났으며, 조석의 영향을 받는 행주 및 김포지점의 관측자료와 팔당호 내에 위치해 있는 팔당 1지점의 경우는 계절적 변동특성이 다소 약한 것으로 나타났다.

4. 결론

한강수계내 팔당댐 이하에 위치한 11개 수질관측지점에서 관측된 계절성(주기성) 및 결측자료를 포함하는 수질 시계열자료를 분석하기 위해서 계절적 Kendall 검정기법과 계절적 Kendall 변화량 추정기법 및 동질성분석을 실시한 결과 다음과 같은 결론을 얻었다.

1. 인접한 관측자료간의 동질성을 분석한 결과 DO, BOD, COD의 경우 보광-노량진과 가양-행주지점에서 추이의 방향이 변화하는 것으로 나타났다.
2. 추이의 존재여부 및 추이의 크기를 분석해 본 결과 하류로 이동할수록 전반적으로 수질은 악화되지만, 본 연구 대상구간의 중상류부에서는 수질이 다소 악화되는 추이를 보이고 있으며, 중하류 이하에서는 수질변화의 추이는 양호하게 개선되고 있는 것으로 나타났다.
3. 조석의 영향을 받는 행주 및 김포지점 이외에서는 수질자료가 현저한 계절적 변동특성을 보이는 것으로 나타났다.

한편, 하천 수질은 유량 및 유역의 토지이용 및 생활환경의 변화, 지류의 유입 등에 밀접한 관계가 있으므로 추후 수량 및 유역의 변화를 고려한 수질의 추이를 분석하고자 한다.

참고문헌

- Gilbert, R. O. (1987). *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinhold, New York.
- Hipel, K. W., McLeod, A. I., and Weiler, R. R. (1988). "Data analysis of water quality time series in lake Erie." *Water Resources Bulletin*, 24(3), pp. 533-544.
- Hirsch, R. M. and Slack, J. R. (1984). "A nonparametric trend test for seasonal data with serial dependence." *Water Resour. Res.*, 20(6), pp. 727-732.
- Hirsch, R. M., and Slack, J. R., and Smith, R. A. (1982). "Techniques of trend analysis for monthly water quality data." *Water Resour. Res.*, 18(1), pp. 107-121.
- Kendall, M. G. (1975). *Rank Correlation Methods*, Charles Griffin, London.
- Lettenmaier, D. P., Hooper, E. R., Wagoner, C., and Faris, K. B. (1991). "Trends in stream quality in the continental United States, 1978-1987." *Water Resour. Res.*, 27(3), pp. 327-339.
- Lettenmaier, D. P. (1976). "Detection of trends in water quality data from records with dependent observations." *Water Resour. Res.*, 12(5), pp. 1037-1046.
- Mann, H. B. (1945). "Non-parametric tests against trend." *Econometrica*, 13, pp. 245-259.
- Pearson, E. S., and Hartley, H. O. (1966). *Biometrika Tables for Statisticians*, Vol 1, 3rd ed. Cambridge University Press, London.
- Sen, P. K. (1968). "Estimates of the regression coefficient based on Kendall's tau." *J. Am. Statist. Assoc.*, 63, pp. 1379-1389.
- Walker, W. W. (1991). "Water quality trends at inflows to everglades national park." *Water Resources Bulletin*, 27(1), pp. 59-72.
- 환경처. (1988-1994). 환경연감.

표1 관측지점별 수질항목에 대한 추이분석 및 관측지점간 동질성 분석결과

수질항목	관측소	자료수	평균	중간값	표준편차	해곡도계수	Seasonal Kendall	Prob.	LL	변화율 (Unit/y)	UL	P(χ^2)
DO	팔당1	104	10.328	10.000	1.834	.2898	-2.6594	.008	-.200	-.094	.000	.546
	팔당2	105	10.430	10.200	1.877	.2988	-1.8633	.062	-.139	-.086	.019	.557
	구리	106	10.459	10.200	1.841	.3893	-.9304	.352	-.108	-.033	.064	.523
	구의	108	10.361	10.200	1.846	.2869	-.0303	.976	-.075	.000	.100	.729
	독도	107	9.975	9.700	1.847	.4223	.3985	.690	-.100	.015	.100	.286
	보광	107	9.507	9.000	1.989	.2331	-.9793	.328	-.167	-.067	.100	.014
	노랑진	108	8.886	8.700	1.903	.1114	2.3680	.018	.000	.078	.207	.617
	영등포	108	8.067	8.300	2.343	-.3741	1.6583	.097	-.100	.142	.324	.147
	가양	108	6.095	6.700	3.104	-.2963	3.7086	.000	.200	.571	.864	.060
	행주	107	6.034	6.400	2.713	-.1672	1.0082	.313	-.200	.140	.436	.627
	김포	63	5.562	5.800	2.326	-.2114	.36447	.715	-.500	.112	.576	
BOD	팔당1	104	1.403	1.400	.396	.4308	-1.3460	.178	-.075	-.018	.029	.973
	팔당2	105	1.186	1.200	.295	.7383	-1.3428	.179	-.050	-.013	.000	.003
	구리	106	1.547	1.500	.495	.6483	2.7126	.007	.000	.050	.100	.675
	구의	108	1.681	1.500	.520	.5317	2.0250	.043	.000	.031	.080	.249
	독도	107	2.697	2.600	1.002	.9658	.4008	.689	-.050	.000	.100	.647
	보광	107	3.319	3.000	1.314	1.3063	1.0713	.284	-.067	.046	.138	.006
	노랑진	108	3.835	3.600	1.537	1.0347	-2.8203	.005	-.221	-.115	.000	.111
	영등포	108	4.431	4.200	2.167	2.0085	-5.0739	.000	-.400	-.225	-.122	.058
	가양	108	7.471	5.800	5.805	4.1193	-7.7583	.000	-1.167	-.900	-.618	.017
	행주	108	7.594	6.800	4.622	4.9937	-4.3799	.000	-.556	-.280	-.100	.660
	김포	68	12.299	7.300	19.044	4.5004	-3.6953	.000	-1.467	-.800	-.299	
COD	팔당1	104	2.137	2.000	.656	2.6303	1.8736	.061	.000	.031	.067	.000
	팔당2	105	2.284	2.100	.872	1.8895	-4.6271	.000	-.200	-.114	-.050	.002
	구리	106	2.806	2.700	.997	2.2606	-.3425	.732	-.100	.000	.067	.591
	구의	108	2.944	2.800	.900	1.0101	-1.1533	.249	-.132	-.050	.055	.978
	독도	107	3.984	3.800	1.259	.7976	-1.0438	.297	-.100	-.039	.050	.035
	보광	107	4.514	4.400	1.409	.3500	1.7784	.075	-.0580	.080	.200	.022
	노랑진	108	4.885	4.900	1.411	.3651	-1.3676	.171	-.154	-.067	.038	.086
	영등포	108	5.275	5.100	2.060	2.3103	-3.7838	.000	-.325	-.185	-.060	.353
	가양	108	7.341	6.500	4.663	3.7950	-5.0955	.000	-.757	-.541	-.262	.006
	행주	108	7.721	7.100	7.770	4.6136	-1.2403	.215	-.364	-.137	.123	.162
	김포	67	11.248	7.900	13.139	4.2601	-3.2778	.001	-1.103	-.550	-.095	
SS	팔당1	104	5.473	4.100	5.797	4.9879	-1.3949	.163	-.450	-.180	.100	.086
	팔당2	105	3.758	3.000	2.391	2.6515	.97448	.330	-.100	.050	.150	.498
	구리	106	6.891	5.100	6.238	4.6621	.06201	.951	-.350	.000	.333	.110
	구의	108	7.071	6.000	4.225	1.4244	-2.1812	.029	-.697	-.300	.009	.143
	독도	107	10.167	8.000	7.021	2.0138	-4.2793	.000	-1.266	-.850	-.324	.609
	보광	106	10.797	8.200	7.161	1.5036	-3.5322	.000	-1.451	-.900	-.221	.836
	노랑진	108	11.308	8.000	8.038	1.6972	-3.8257	.000	-1.605	-.917	-.170	.526
	영등포	108	13.893	10.100	11.967	3.6415	-4.7294	.000	-1.988	-1.200	-.570	.501
	가양	108	15.024	10.000	12.491	2.4342	-3.7758	.000	-2.532	-1.275	-.368	.560
	행주	108	20.600	14.600	19.649	3.3181	-2.9535	.003	-2.600	-1.500	-.130	.855
	김포	67	31.664	18.000	35.553	11.8399	-3.2879	.001	-7.227	-2.850	-.729	
pH	팔당1	104	7.711	7.700	.387	.0589	-3.3352	.001	-.100	-.057	.000	.209
	팔당2	105	7.554	7.500	.294	.1329	-1.5573	.119	-.050	-.017	.007	.001
	구리	106	7.508	7.500	.306	.0783	2.9720	.003	.000	.025	.055	.965
	구의	108	7.511	7.500	.304	.1237	3.0436	.002	.000	.033	.067	.609
	독도	107	7.479	7.500	.311	.1932	2.2607	.024	.000	.025	.050	.422
	보광	107	7.426	7.400	.275	.3649	1.1838	.237	-.014	.000	.036	.606
	노랑진	108	7.416	7.500	.247	.1578	1.9204	.055	.000	.014	.037	.372
	영등포	108	7.360	7.400	.291	-.0747	3.1878	.001	.000	.038	.067	.504
	가양	108	7.357	7.300	.340	1.3608	4.1504	.000	.017	.050	.075	.025
	행주	108	7.329	7.300	.322	1.0797	.9525	.341	-.017	.000	.040	.804
	김포	67	7.303	7.200	.272	.5848	.75223	.452	-.027	.000	.041	

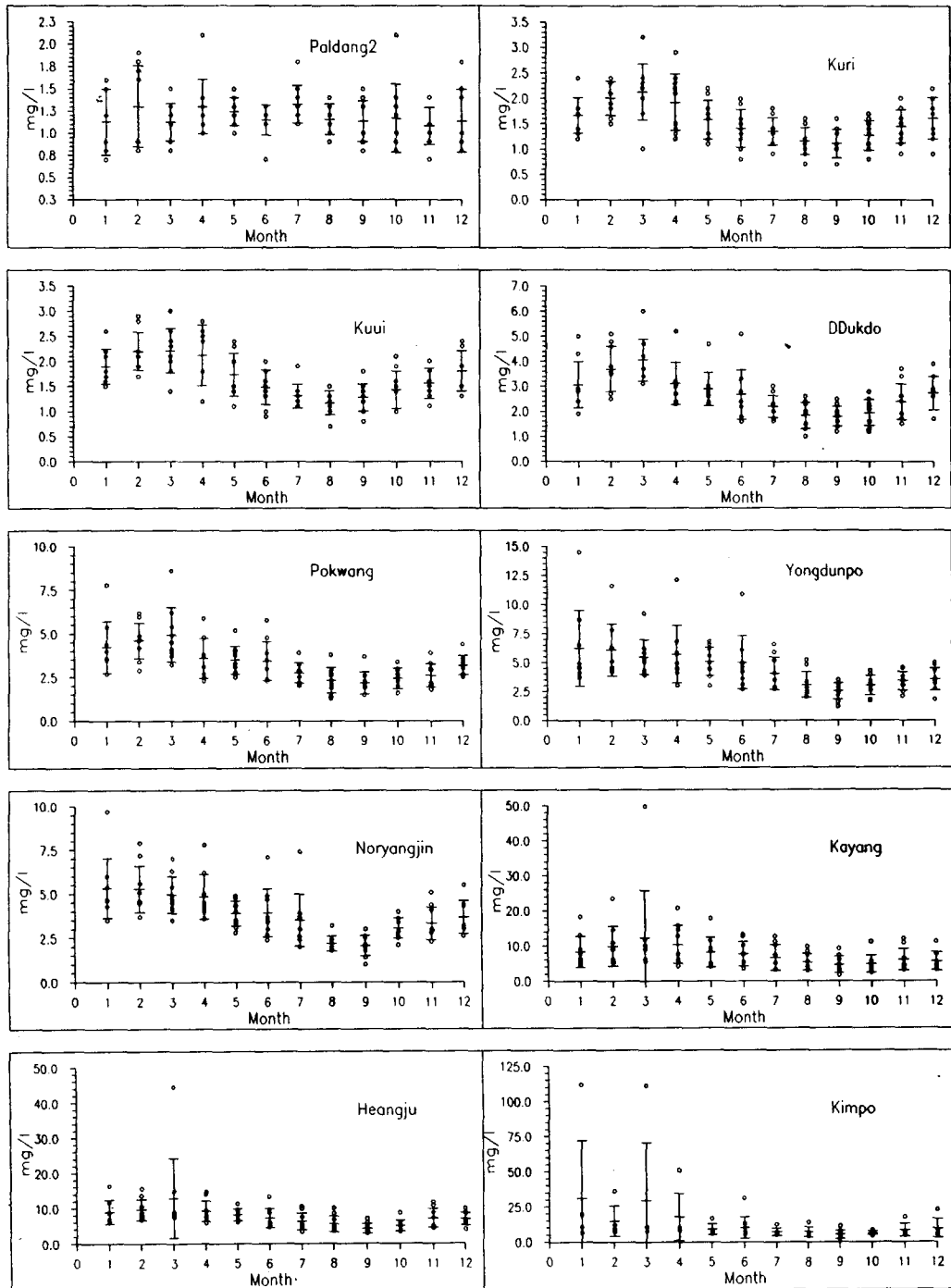


그림 1 월별 BOD 자료의 분포도