

## Forecasting of Stream Qualities in Gumho River by Exponential Smoothing at Gumho2 Measurement Point using Monthly Time Series Data<sup>1)</sup>

Phil Jun Song<sup>2)</sup> · Bora Lee<sup>3)</sup> · Jin Yong Kim<sup>4)</sup> · Jong tae Kim<sup>5)</sup>

### Abstract

The goal of this study is to forecast the trend of stream quality and to suggest some policy alternatives in Gumho river. It used the five different monthly time series data such as BOD, COD, T-N and EC of the nine of Gumho River measurement points from Jan. 1998 to Dec. 2006. Water pollution is serious at Gumho2 and Palgeo stream measurement points. BOD, COD, T-N and EC data are analyzed with the exponential smoothing model and the trend is forecasted until Dec. 2009.

**Keywords** : Exponential Smoothing, Time Series Anaysis, Water Quality Forecasting

### 1. 서론

낙동강 수질 오염에 크게 영향을 주고 있는 지천 중의 하나가 금호강이라 보고 연구대상 지역으로 선정하게 되었다. 본 연구는 수질측정자료를 이용하여 금호강유역 9개 지역을 대상으로 1998년 1월부터 2006년 12월까지의 BOD, COD, 총질소 및 전기전도도 변화 경향을 파악하였다.

금호강1 배수구역은 전부가 경상북도에 위치하고 있으며 영천시 전부와 경산시, 포항시의 일부를 포함한다. 금호강2 배수구역은 경상북도 경산시 일부, 대구광역시의 동구 전부와 수성구와 북구 일부를 포함하고 있다. 신천 다음으로 오는 지역인 팔계천

---

1) 이 논문은 2006년도 대구대학교 학술연구비 지원에 의한 연구임

2) Professor, Department of Computer Science & Statistics Daegu University, Kyungbook 712-714

3) Student, Department of Computer Science & Statistics Daegu University, Kyungbook 712-714

4) Student, Korean Minjok Leadship Academy, 1334, Sosa, Anhung, Gangwan 225-823

5) Professor, Department of Computer Science & Statistics Daegu University, Kyungbook 712-714 Korea. E-mail : jtkim@daegu.ac.kr

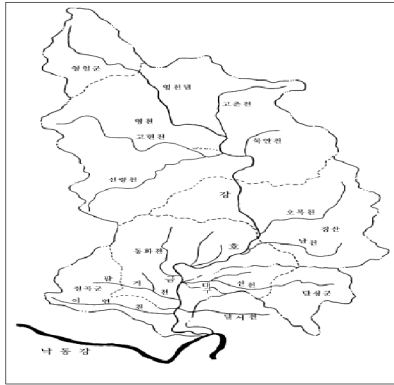
배수구역은 전부가 대구광역시에 위치하고 있다. 금호강2 배수구역은 대구광역시 일부를 포함하고 있으므로 생활하수의 유입이 금호강1 배수구역보다 많다. 팔계천 배수구역은 많은 공단이 위치하고 있으므로 산업폐수에 의한 수질오염 실태를 반영하고 있다. (김승우 외 1명(1996), Zun-Ung at. al.(2002)). 수질오염에 크게 기여하는 금호강2 지역을 선별하여 2009년 12월까지의 추세 변화 예측을 시도한 것이다. (낙동강유역환경청(2003)).

### 2. 연구방법 및 연구대상지 선정

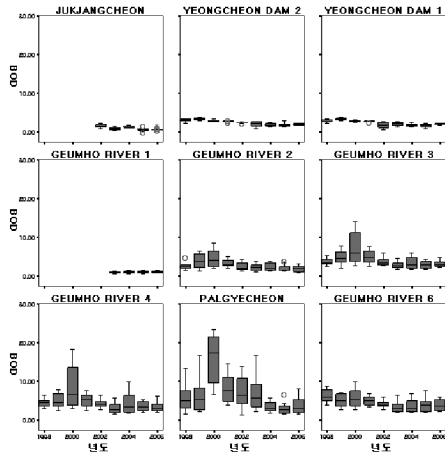
일반적인 통계기법은 관측치의 정상분포와 독립성에 기초를 둔 반면 시계열분석은 관측치의 발생 순서를 보전하는 시계열을 중심으로 관측치의 자기상관성에 기초를 두고 있다. 특정변수의 시간에 따른 변화양상을 파악하기 위한 시계열분석은 각 관측치

간의 상호관계를 밝히고 이를 바탕으로 한 모형을 확정하여 미래를 예측하고자 하는 체계적인 방법이다. (조덕호 외 1명(2005)). 시계열분석을 통한 예측에서는 관측된 과거의 자료를 분석하여 규칙성을 발견하고, 이러한 규칙성이 미래에도 변하지 않고 계속될 것이라는 가정하에 이를 모형화하여 추정하고, 추정된 모형을 이용하여 미래의 시계열을 예측하게 된다. (참조; 이덕기(1999)).

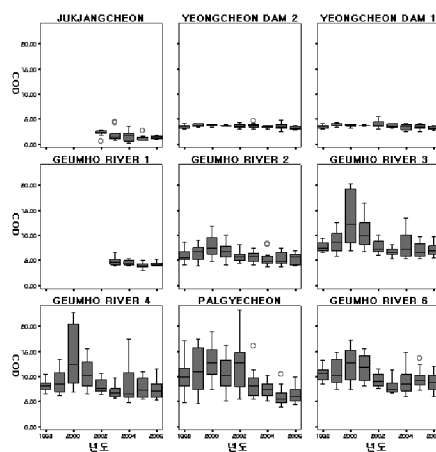
본 연구에 앞서 금호강유역을 상류(죽장천, 영천댐2, 영천댐1), 중류(금호강1, 금호강2, 금호강3), 하류(금호강4, 팔계천, 금호강5)로 분류해 수질 오염에 많은 영향을 미치는 지역을 찾아보았다.



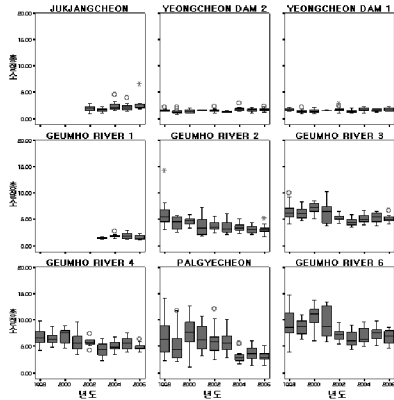
[그림1] 금호강유역의 수계와 행정구역



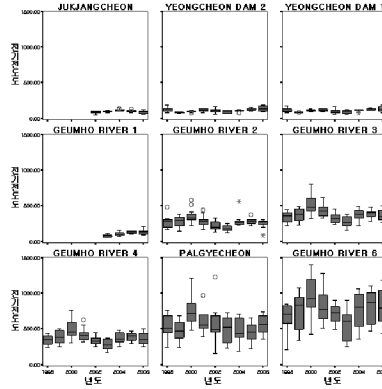
[그림2-1] BOD의 연도별 변화



[그림2-2] COD의 연도별 변화



[그림2-3] 총질소의 연도별 변화

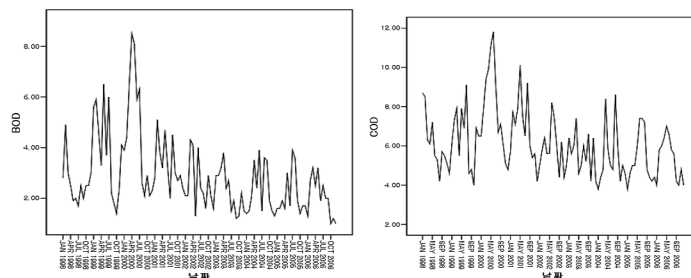


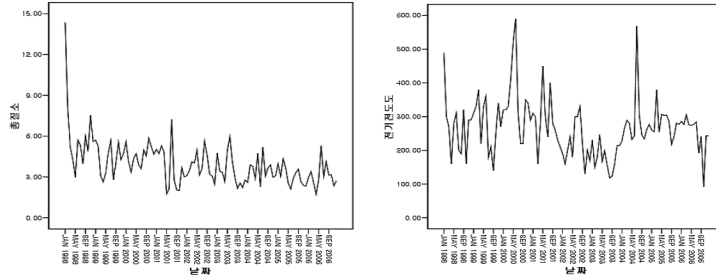
[그림2-4] 전기전도의 연도별 변화

위 상자도표들은 본 연구에 사용된 수질환경오염지표의 1998년 1월부터 2006년 12월까지의 변화를 지역별로 나타낸 것이다. 상류 지역의 BOD농도는 다른 지역에 비해 낮게 나타났고, BOD의 변화추세를 살펴보면 시간이 지남에 점차적으로 수질이 개선되고 있음을 알 수 있다. 금호강1과 금호강2 지역에 차이가 나는 것으로 보이며 팔계천에서는 다른 지역에 비해 BOD가 높게 측정된 것으로 나타났다. [그림2-1] [그림2-2]를 보면 금호강2, 금호강3, 금호강4 지역에 이르기까지 2000년도 관측치가 시간이 지남에 따라 급속히 높게 나타났다. 팔계천은 시간의 경과에 따라 빠른 속도로 개선되고 있고, 금호강2의 경우 미약하지만 조금씩 수질이 개선되고 있음을 알 수 있다. 지역별 총질소 부하량을 비교했을 때 상류에 비해 중류와 하류에서 크게 나타나고 있는 것으로 보아 대구지역의 높은 인구밀도와 하수처리의 부족에 의한 것으로 본다. [그림2-3] 전기전도도는 다른 매개변수들과는 다른 패턴을 보이고 있고 낙동강 유역으로 근접 할수록 높은 수치를 보인다. [그림2-4] 분석결과 금호강2와 팔계천 두 배수구역이 수질 오염에 많은 영향을 주는 것으로 추정된다.

### 3. 수질예측모형의 선택, 분석

다음의 [그림3-1]에서 수질측정자료의 1998년 1월부터 2006년 12월까지의 BOD, COD, 총질소, 전기전도도의 변화를 순차도표로 나타낸 것이다.





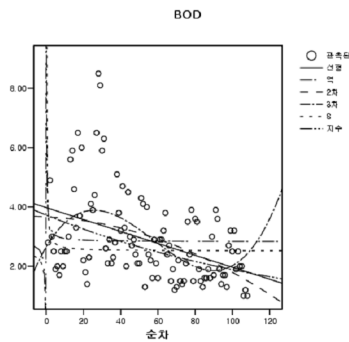
[그림3-1] 금호강2의 수질특성변화

3.1. 모형선택

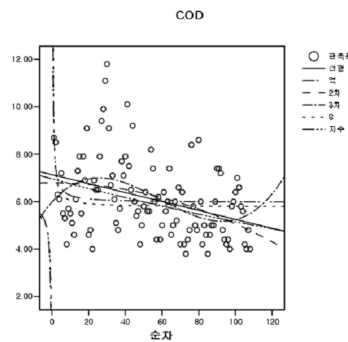
매개변수 변화추세를 검증 및 확인하기 위하여 월별 시계열자료를 비선형추세 시계열 평활법을 사용하여 추세변화 예측을 시도하였다. 비선형추세 시계열 평활법에는 역선형추세모형, 2차모형, 3차모형, S추세모형, 지수곡선모형 등이 있으나 잠정적인 모형들 중 평균자승오차(MSE)를 가지고 적합한 모형을 선택했다.

<표3-1> BOD의 모형선택

| 모형       | 모형식   | 계수추정  | 평균자승오차    |
|----------|---|---|-----------|
| LINEAR   | $Y_t = a + bt + \epsilon_t$                     | $a = 3.977155 \quad b = -.020121$   | 1.841815  |
| INVERSE  | $Y_t = a + b/t + \epsilon_t$                    | $a = 2.828565 \quad b = 1.066678$   | 2.2279242 |
| QUADRATI | $Y_t = a + b_1t + b_2t^2 + \epsilon_t$          | $a = 3.66756 \quad b_1 = -.00323$<br>$b_2 = -.000155$                       | 1.840705  |
| CUBIC    | $Y_t = a + b_1t + b_2t^2 + b_3t^3 + \epsilon_t$ | $a = 2.648542 \quad b_1 = .106432$<br>$b_2 = -.00266 \quad b_3 = .5313E-05$ | 1.720558  |
| S        | $Y_t = e^{a+b/t} + \epsilon_t$                  | $a = .921682 \quad b = .470280$   | .21857164 |
| EXPONENT | $Y_t = ae^{bt} + \epsilon_t$                    | $a = 3.726027 \quad b = -.006803$   | .1756256  |



[그림3-2] BOD의 수질특성변화



[그림3-3] COD의 수질특성변화

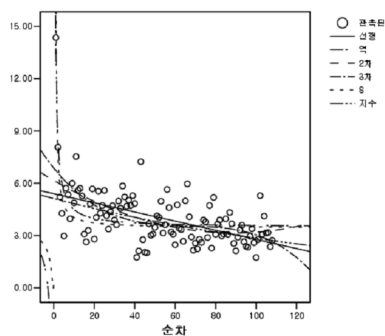
<표3-2> COD의 모형선택

| 모형       | 모형식   | 계수추정  | 평균자승오차    |
|----------|---|---|-----------|
| LINEAR   | $Y_t = a + bt + \epsilon_t$                     | $a = 7.1379 \quad b = -.018839$   | 2.418793  |
| INVERSE  | $Y_t = a + b/t + \epsilon_t$                    | $a = 5.9641 \quad b = 3.016271$   | 2.651880  |
| QUADRATI | $Y_t = a + b_1t + b_2t^2 + \epsilon_t$          | $a = 6.7935 \quad b_1 = 5.602E-05$<br>$b_2 = -.000172$                      | 2.418753  |
| CUBIC    | $Y_t = a + b_1t + b_2t^2 + b_3t^3 + \epsilon_t$ | $a = 5.984377 \quad b_1 = .087021$<br>$b_2 = -.00216 \quad b_3 = 1.216E-05$ | 2.355104  |
| S        | $Y_t = e^{a+b/t} + \epsilon_t$                  | $a = 1.753310 \quad b = .476622$  | .06342604 |
| EXPONENT | $Y_t = ae^{bt} + \epsilon_t$                    | $a = 6.959847 \quad b = -.003002$   | .05745668 |

<표3-3> 총질소의 모형선택

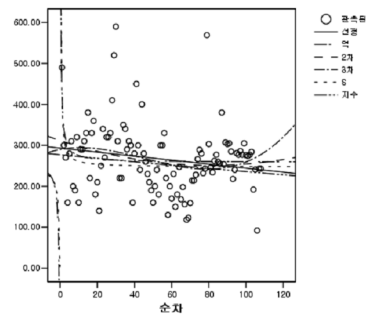
| 모형       | 모형식   | 계수추정  | 평균자승오차   |
|----------|---|---|----------|
| LINEAR   | $Y_t = a + bt + \epsilon_t$                     | $a = 5.397734 \quad b = -.026208$   | 2.005439 |
| INVERSE  | $Y_t = a + b/t + \epsilon_t$                    | $a = 3.474441 \quad b = 10.155136$  | 1.34383  |
| QUADRATI | $Y_t = a + b_1t + b_2t^2 + \epsilon_t$          | $a = 6.157706 \quad b_1 = -.067661$<br>$b_2 = .000380$                        | 1.912147 |
| CUBIC    | $Y_t = a + b_1t + b_2t^2 + b_3t^3 + \epsilon_t$ | $a = 6.867381 \quad b_1 = -.144036$<br>$b_2 = .002124 \quad b_3 = -1.066E-05$ | 1.863676 |
| S        | $Y_t = e^{a+b/t} + \epsilon_t$                  | $a = 1.235007 \quad b_1 = 1.592854$   | .0939461 |
| EXPONENT | $Y_t = ae^{bt} + \epsilon_t$                    | $a = 5.12090 \quad b_1 = -.005884$  | .0926711 |

총질소



[그림3-4] 총질소의 수질특성변화

전기전도도



[그림3-5] 전기전도도의 수질특성변화

&lt;표3-4&gt; 전기전도도의 모형선택

| 모형       | 모형식   | 계수추정   | 평균자승오차    |
|----------|---|--|-----------|
| LINEAR   | $Y_t = a + bt + \epsilon_t$                     | $a = 293.162340$ $b = -.489387$                                      | 7186.749  |
| INVERSE  | $Y_t = a + b/t + \epsilon_t$                    | $a = 258.260624$ $b = 168.855969$                                    | 7052.945  |
| QUADRATI | $Y_t = a + b_1t + b_2t^2 + \epsilon_t$          | $a = 311.684932$ $b_1 = -1.499710$<br>$b_2 = .009269$                | 7188.431  |
| CUBIC    | $Y_t = a + b_1t + b_2t^2 + b_3t^3 + \epsilon_t$ | $a = 289.715825$ $b_1 = .864599$<br>$b_2 = -.044709$ $b_3 = .000330$ | 7193.481  |
| S        | $Y_t = e^{a+b/t} + \epsilon_t$                  | $a = 5.509934$ $b_1 = .532174$                                       | .09919025 |
| EXPONENT | $Y_t = ae^{bt} + \epsilon_t$                    | $a = 277.003818$ $b_1 = -.001618$                                    | .10028395 |

잠정적으로 설정된 모형들의 MSE를 보면 가장 작은 MSE를 보이고 있는 지수모형이 원시 자료를 잘 적합 시킨다고 보고 지수곡선모형으로 추세변화를 분석하였다. BOD의 경우  $Y_t = 3.726027e^{-.006803t}$ , MSE=.1756256 로 시간이 지남에 따라 조금씩 개선되고 있다고 본다. COD의 경우에는  $Y_t = 6.959847e^{-.003002t}$ , MSE=.05745668, 총질소는  $Y_t = 5.120900e^{-.005884t}$ , MSE=.0926711로 나타나 큰 변화는 보이지 않지만 화학적 산소 요구량과 총질소 모두 줄어 들고 있다. 전기전도도의 경우  $Y_t = 277.003818e^{-.001618t}$ , MSE=.10028395로 빠른 속도로 변하는 추세다.

### 3.2. 연구대상지의 수질 예측

분석한 자료의 특성으로 볼 때 1년(12개월) 단위로 계절변동주기를 가지고 있는 것으로 볼 수 있으며, 매개변수의 관측치 들이 시간이 지남에 따라 낮게 나타나는 것으로 보아 수질이 점차 개선되고 있다고 예측할 수 있다. 따라서 추세와 계절유형을 동시에 보이는 이 데이터를 지수평활법으로 모형화 하고 윈터스의 지수평활법을 사용하여 미래시계열을 예측하였다. 윈터스의 지수평활법의 기본 모형식은 다음과 같다.

$$\text{수준모형 : } L_t = \alpha \left( \frac{Y_t}{S_{t-m}} \right) + (1-\alpha)(L_{t-1} + T_{t-1})$$

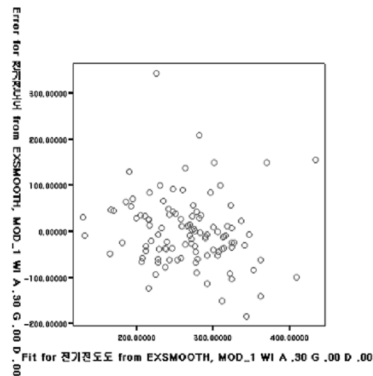
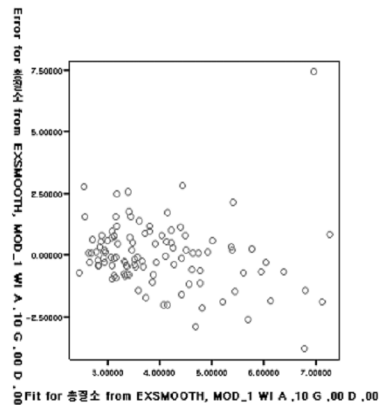
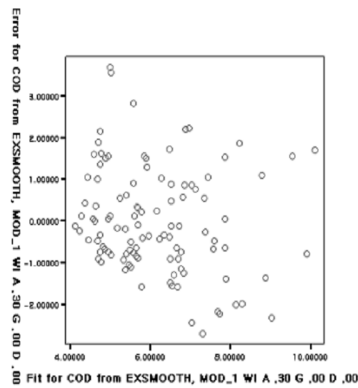
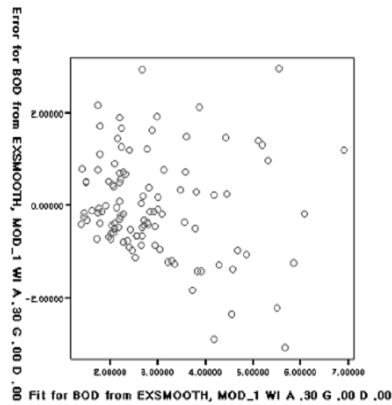
$$\text{추세영향치 : } T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

$$\text{계절영향치 : } S_t = \gamma \left( \frac{Y_t}{L_t} \right) + (1-\gamma)S_{t-m}$$

$$\text{예측치 : } F_{t+k} = (L_t + kT_t)S_{t+k-m}$$

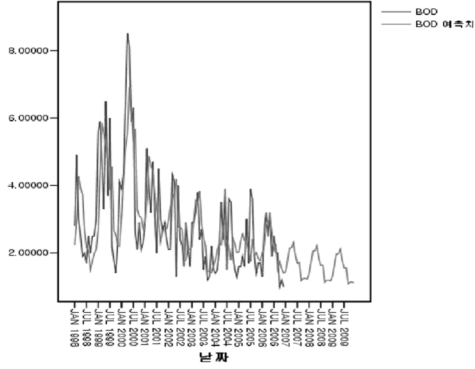
<표3-5>가장 작은 제곱오차를 제공하는 평활상수 값

|       | $\alpha$ | $\beta$ | $\gamma$ |
|-------|----------|---------|----------|
| BOD   | 0.3      | 0       | 0        |
| COD   | 0.3      | 0       | 0        |
| 총질소   | 0.1      | 0       | 0        |
| 전기전도도 | 0.3      | 0       | 0        |

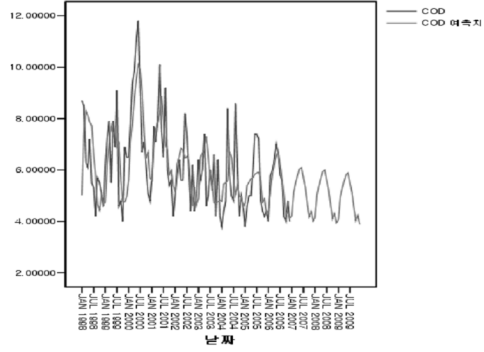


[그림3-6] 금호강2 잔차를 통한 모형적합도 검진

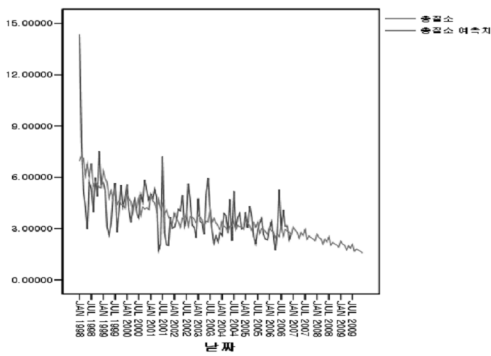
[그림 3-6]를 분석한 결과 잔차들이 평균 0을 중심으로 특정한 패턴 없이 무작위로 존재하기 때문에 추정된 모형은 적절하다고 보고 이 모형을 이용하여 미래를 예측하면 다음과 같이 나타났다.



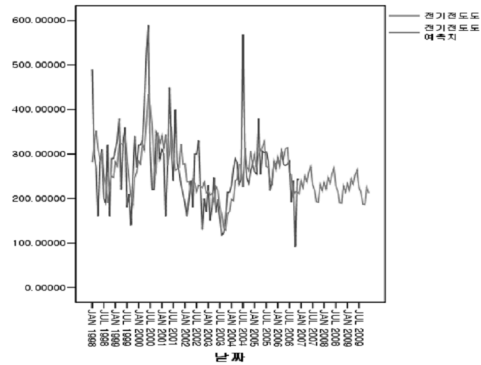
[그림3-7] BOD 모형적합도 검진



[그림3-8] COD 모형적합도 검진



[그림3-9] 총질소 모형적합도 검진



[그림3-10] 전기전도도 모형적합도 검진

<표3-6> 금호강2의 관측치와 예측치간의 통계량 비교

|           | N   | 평균     |         | 표준편차     | 분산       |
|-----------|-----|--------|---------|----------|----------|
|           | 통계량 | 통계량    | 표준오차    | 통계량      | 통계량      |
| BOD 관측치   | 108 | 2.88   | .14343  | 1.49056  | 2.222    |
| BOD 예측치   | 144 | 2.58   | .09845  | 1.18137  | 1.396    |
| COD 관측치   | 108 | 6.11   | .15941  | 1.65661  | 2.744    |
| COD 예측치   | 144 | 5.85   | .10777  | 1.29328  | 1.673    |
| 총질소 관측치   | 108 | 3.97   | .15695  | 1.63110  | 2.660    |
| 총질소 예측치   | 144 | 3.54   | .10342  | 1.24107  | 1.540    |
| 전기전도도 관측치 | 108 | 266.49 | 8.25212 | 85.75857 | 7354.533 |
| 전기전도도 예측치 | 144 | 256.82 | 4.29280 | 51.51364 | 2653.655 |

BOD농도를 분석한 결과 관측값과 예측값의 차트는 [그림3-7]과 같다. BOD의 예측값은 큰 변화를 보이지 않고 시간이 지남에 따라 점점 낮게 측정될 것으로 본다. 조덕호 외 1명(2005) 낙동강 하천수질예측에 관한 연구에서 ARIMA 모형에 적합시킨



것과는 다소 차이를 보이고 있고 다른 매개변수에서 얻은 차트와는 달리 일치도가 우수하다고는 할 수 없지만 기술통계량을 비교해 보았을 때 평균과 표준편차에 큰 차이가 없는 것으로 보아 적합한 모형이라 본다. (조덕호 외 1명 (2004), 참조).

COD의 예측값은 계절성을 띄고 있으며 2007년부터 점차적으로 줄어 들 것으로 예측 할 수 있다. 기술통계량을 기준으로 관측치와 예측치 사이에 큰 차이를 보고 있지 않기 때문에 모형이 적절한 것으로 보인다.

#### 4. 결론

본 연구에 따르면 2000년도 이후에 시행된 수질관리정책을 통해 차츰 수질이 개선되고 있다고 판단했다. 그러나 금호강 지류의 다른 지역에 비해 금호강2의 경우 낙동강 수질에 큰 영향을 주는 것으로 보고 이 지역을 중심으로 BOD, COD, 총질소 및 전기전도도의 데이터를 매개변수로 지정하여 수질을 분석하고 윈터스의 지수평활법으로 2009년 12월까지의 추세를 예측하였다. 시간이 지남에 따라 규제의 강화로 수질이 개선되고는 있으나 향후 더 나은 수질개선을 위해 규제를 강화하고 배출량에 따른 부과금을 징수하도록 하고, 수질총량관리시스템 구축에 힘써야 할 것이다.

#### 참고문헌

1. 김승우, 김희성(1996), 금호강유역에서의 수질총량규제 실시방안 연구, 한국환경기술개발원, 기본과제 연구보고서, KETRI/RE-17.
2. 낙동강유역환경청(2003), 오염관리제도 추진상황, 공개자료.
3. 이덕기(1999), 예측방법의 이해, SPSS 아카데미
4. 조덕호, 엄홍석(2005), 낙동강 하천수질예측에 관한 연구-월별 시계열자료를 이용한 고령지점을 중심으로, 환경정책 제13권, 2호, 5-28.
5. Zun-Ung Bae, Sang-Hak Lee, Seong-Ho Lee(2002), A Survey of the Kumho River Pollution, *Analytical Science Technology*, Vol. 15, 54-66.

[ 2007년 8월 접수, 2007년 8월 채택 ]