

## ARIMA 모델에 의한 上水道 一日 給水量 推定에 관한 研究

### A Study on the Estimation of Daily Urban Water Demand by ARIMA Model

이 경 훈\* · 문 병 석\*\* · 박 성 천\*\*\*

Rhee, Kyoung-Hoon · Moon, Byoung-Seok · Park, Sung-Chun

---

#### Abstract

The correct estimation of the daily or hourly urban water demand is required for the efficient management and operation of the water supply facilities. The prediction of water supply demand are regression model and time series method, the optimum ARIMA (Auto Regressive Integrated Moving Average) model was sought for the daily urban water demand estimation in this paper. The data used for this study were obtained from the city of Kwangju, Korea. The raw data used in this study were rearranged 15, 30, 60, 90 days for the purpose of analysis. The statistical analysis was applied to the data to obtain the ARIMA model. As a result, the parameters detemining the ARIMA model was obtained. The accuracy of the model was 2% of mean average error and 12% of maximum error, except an accident or suspension of water supply. The developed model was found to be useful for the practical operation and management of the water supply facilities.

---

#### 요 지

수도물, 송배수펌프의 운전 등 상수도시설을 합리적인 운용을 위해서는 日, 또는 時間 單位의 給水量 사용량의 推定이 필수적이라고 할 수 있다. 給水量的 推定方法은 회귀모형식 및 시계열 분석방법이 있는데, 본 연구에서는 시계열 분석방법인 ARIMA 모형을 이용하여 一日 給水량을 推定하였으며 연구대상 지역으로는 광주광역시를 선정하였다. 一日 給水량을 推定하는데 있어서 時系列長을 15, 30, 60, 90일로 나누어 각각의 時系列長에 대해 시행착오법으로 각 모형에 적용하여 最適의 時系列長을 결정하여, 上水道 一日 給水량을 推定할 수 있는 모형을 제안하고 그 有效性을 殘差分析을 통해 檢證하였다. 제안된 모형식은 사고 등의 人爲的인 操作(斷水 등)이 가해지는 시기를 제외하고는 실측치와 모형의 추정치와의 誤差率이 최대 약 12%, 평균 3% 이내로 나타나, 모형의 결과는 上水道 一日 給水量的 推定에 필요한 시설에 적용 가능하다고 판단된다.

---

\* 전남대학교 공과대학 토목공학과 부교수  
\*\* 서남대학교 공과대학 토목공학과 전임강사  
\*\*\* 동신대학교 공과대학 환경공학과 전임강사

## 1. 序 論

上水道 1日 給水量을 豫測하는 방식은 과거의 資料로부터 統計學의 方法에 의하여 豫測할 수 있는 多重回歸分析法과 時系列分析法으로 접근할 수 있다. 또한 給水量 豫測은 시점에 따라 短期豫測과 長期豫測으로 분류할 수 있다. 短期豫測은 豫測期間을 時間單位나 日單位로 구분하여 給水量을 推定하며, 豫測된 給水量은 給·配水量을 실시간으로 제어하는 데 사용한다. 長期豫測은 豫測구간을 月別, 季節別 또는 年別로 구분하여 給水量을 豫測하며, 이는 上水道施設을 計劃하거나 設計하는데 필요한 資料로 활용하고 있다. 하지만 지금까지의 研究는 주로 上水道施設을 計劃하거나 設計하는데 필요한 長期豫測 위주로 研究가 진행되어 給·配水量 制御에 필요한 短期豫測에 대한 연구는 미흡한 실정이다.

給·配水量을 制御하기 위해서는 日別 또는 時間別 데이터에 의한 上水 使用量의 確率的 豫測이 필요하다. 日 또는 몇 時間前의 上水資料에 의한 上水道 運營計劃을 세우게 되면, 合理的인 送配水 펌프의 운영, 밸브의 조정 및 配水池 水位의 설정 등이 용이하게 될 것이다.

給水量 豫測에 관한 연구동향을 보면 Young (1973)은 年別로 변화되는 給水量의 資料를 해당 도시의 인구수, 세대의 평균수입, 수도요금과 더불어 강우량, 기온, 증발량을 연관시키려고 시도하였으며, Frankel과 Shouvanavirakul(1973)은 季節別 給水需要 模型을 1차선형 多重回歸分析 方法으로 계산하였다. Maidment와 Parzen(1984)은 月別 平均給水量을 강우량, 증발량, 기온 등 氣候因子를 사용하여 1日 給水量을 豫測한 바 있다. 比企三藏(1990)은 給水量의 변화를 기후예보 데이터 및 전일의 配水量을 변환팩타로 만들어 Kalman Filter에 의해 1日 給水量을 豫測한 바 있으며, 小泉明 등(1988)은 上水使用量에 영향을 주는 因子를 月別로 선정하여 多元 ARIMA모형으로 1日 給水量을 豫測한 바 있다.

國內의 연구동향을 보면, 김수원(1985)은 都市給水에 영향을 주는 要因을 說明變數로 하여 用途別로 需要豫測을 하였고, 이경훈 등(1994a, b, c)은 上水道 1日 給水量을 時間別, 曜日別, 季節의 분포에 관

한 연구를 하였으며, 현인환과 목동우(1995), 엄동조(1995)는 給水量에 영향을 주는 因子로 景年 변화, 인구수, 최고기온, 강우량, 운량, 일조시간, 일사량 등을 사용하여 多重回歸分析에 의한 1日 給水量을 豫測한 바 있다. 이들 1日 給水量 豫測에 관한 연구는 기후인자 등의 자료를 이용하여 급수량을 예측에 필요한 매개변수의 추정에 지나지 않으며, 실제 추정된 매개변수를 이용하여 익일 등의 급수량을 예측한 것은 아니다.

따라서 本 研究에서는 翌日의 給水量을 推定하기 위해 日給水量 資料를 時間從屬의 模型인 ARIMA모형을 이용하였으며 대상지역 광주광역시를 선정하였다. ARIMA모형으로 給水量을 推定하기 위해서는 먼저 資料의 定常性을 검토하여야 한다. 따라서 本 研究에서는 資料의 定常性을 자기상관함수로 定常性을 검증하였고, 非定常的인 時系列을 次分하여 非定常的인 時系列 資料를 定常時系列 資料로 변환하였다.

또한 ARIMA모형의 次數(p, q)를 결정하기 위해 次數를 3차 까지 조합하여 모형을 구성하여 각 모형에 대해 10% 이상의 誤差가 발생하거나, 수렴하지 않은 모델은 제외시켰다. 그리고 時系列長 15, 30, 60, 90일을 구성된 각 모델에 대해 적용하여 時系列長을 결정하였다.

## 2. 對象地域 및 研究資料

本 研究의 대상지역으로는 光州廣域市로 선정하고, 給水量資料를 확보하여 本 研究에 활용하였다. 光州廣域市の 1994년 12월말 현재 人口는 1,274,000 여명에 이르고 있으며, 面積은 약 500km<sup>2</sup>로서, 上水使用量이 생활용수를 중심으로 매년 증가하고 있는 지역이다.

光州廣域市の 給水量을 살펴보면 1994년 말 현재 平均 38만톤/日 규모의 給水量에 달하고 있으며, 用途別 給水量은 Table 1과 같다. 광주광역시 용도별 給水量에서 찾아볼 수 있는 특징은 工業用水가 차지하는 比率이 1994년 현재 약 4.32%로서 타 도시에 비하여 公業용수가 차지하는 비중이 타 用途에 비하여 낮은 것이 특징이다. 本 研究에 사용된 資料는 1994년 1월 1일 ~12월 31일 까지의 光州廣域市 上

水道 給水量이다. Fig. 1은 본 논문에서 사용한 1994년도 광주광역시 給水量의 변화를 보여주고 있다.

Table 1. Water Use Classified according to the Using Purpose in the City of Kwangju for the Year of 1994

Usage Year	Domestic use	Industrial use	Commer- cial use	Public & etc.
1989	68.65	4.17	22.65	4.53
1990	69.57	3.87	21.78	4.80
1991	69.04	4.17	21.52	5.27
1992	66.92	5.68	22.02	5.38
1993	68.45	3.75	22.54	5.36
1994	69.78	4.32	21.12	4.78

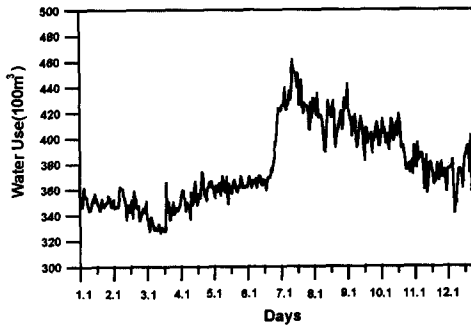


Fig. 1. Variation of Daily Water Use in the City of Kwangju for the Year of 1994

### 3. 上水道 給水量의 變動

#### 3.1 上水道 給水量의 時系列 變動

상수도 給水量은 인구수, 도시의 산업구조, 생활 수준에 따라 변하며 시간별, 일별, 월별 계절별, 또는 년별로 변해가는 특성을 가지고 있다. 이렇게 복잡한 변화 양상을 가지고 있는 상수도 給水量은 시간에 따라 주기적으로 또는 불규칙적으로 변동을 하는 특성으로 인하여 時系列 資料로 분류되고 또한 해석될 수 있다. 時系列 資料는 시간에 따라 변동하기 때문에 이를 분석하거나 변동하는 상태를 推定하기 위하여 時系列 資料의 변동에 영향을 주는 여러 요인으로 분해하

여 이해하려고 하는 시도가 이루어 지고 있다. 時系列 資料를 분석하기 위한 일반적인 방법은 時系列을 趨勢要因, 循環要因, 季節 그리고 不規則要因으로 나누어 分析하는 方法이다.

趨勢要因은 時系列 資料의 장기적 형태에 영향을 주고, 循環要因은 경제 혹은 특정산업에 있어서 나타나는 주기적 변화를 나타낸다. 季節要因은 요일로 반복되거나, 일년중의 각 월에 의한 변화, 사분기 資料에서 각 분기에 의한 변화 등 고정된 주기에 의한 변동을 나타낸다. 季節要因이 年, 月 또는 週期와 같이 설명할 수 있는 고정기간을 반복하는 것인데 반해 循環要因은 季節要因에 비해 비교적 긴 주기를 가지면서 주기가 고정될 명백한 자연적인 이유 등이 없는 要因이다. 不規則要因은 이 세가지 要因으로 설명할 수 없는 誤差에 해당하는 부분이다.

上水給水量의 경우 時系列變動은 연별 증가 또는 감소하는 경우, 봄, 여름, 가을, 겨울에 의해 季節적으로 변화하는 季節變動과 일의 주야로 변화하는 時間變動 등 週期變動, 또한 斷水나 火災 등에 따른 不規則變動 등이 복잡하게 조합되어 있다고 말할 수 있다.

本 研究에서는 一日 上水給水量 變化를 推定하는 것을 目的으로 하고 있기 때문에 趨勢變動, 季節變動은 고려하지 않고 短期 期間範圍(時系列長)을 설정해서 上水給水量을 推定하였다.

#### 3.2 時系列의 定常性

時系列 分析에서 중요한 개념의 하나는 定常性이다. 이를 推計學的 確率過程論에서 도입된 것으로 다음과 같이 정의된다. 어떤 變數列  $\{\dots, Z_{-1}, Z_0, Z_1, \dots\}$ 이 定常性을 갖기 위한 必要充分條件은 임의의 자연수  $m, k$ 와 임의의 시점들  $t_1, t_2, \dots, t_k$ 에 대한  $Z(t_1), Z(t_2), \dots, Z(t_k)$ 와  $Z(t_1 + m), Z(t_2 + m), \dots, Z(t_k + m)$ 이 같은 結合確率分布를 갖는 것으로서 임의의 시점에 대한 時系列의 평균값이 변화되지 않음을 의미하고 있다.<sup>11)</sup> 時系列  $\{Z_t\}$ 가 定常性을 갖으면 다음과 같은 사항을 만족하는 것으로 알려지고 있다.

- (1) 時系列의 平均값  $E(Z_t)$ 는 일정하며, 평균값은 시점  $t$ 에 의존하지 않는다.
- (2) 時系列의 分散  $\text{Var}(Z_t)$ 도 시점  $t$ 에 의존하지

않는다.

(3) 共分散  $Cov(Z_t, Z_s)$ 는 시간차인  $t_s$ 에 의존하나 시점인  $t$ 와  $s$ 에 의존하지 않는다.

이 세가지 조건을 만족하는 時系列은 넓은 의미의 定常性을 갖고 있다고 말한다.

주어진 資料가 定常的이나 非定常的이나의 판단기준은 일반적으로 時系列 散布度 및 自己相關函數에 따른다. 즉, 時系列이 定常的이면 時差가 증가함에 따라 自己相關函數의 絶對值가 감소하여 0에 收斂하게 되지만, 非定常 時系列의 自己相關函數는 時差가 증가하더라도 매우 완만하게 감소한다. 本 研究에서는 時系列長의 定常性을 自己相關函數 方法에 의해 검토하였다.

時系列  $Z_1, Z_2, \dots, Z_n$ 이 주어졌을 때 自己相關函數는 다음과 같다.

$$r_k = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}, k=0, 1, \dots \quad (1)$$

Fig. 2는 1994년도 光州廣域市의 一日 給水量의 自己相關函數를 나타내고 있다. Fig. 2에서 給水量의 自己相關函數는 시간이 경과함에 따라서 완만하게 감소하고 있으며, 이는 給水量의 時系列長이 非定常 時系列임을 뜻하고 있다. 非定常 時系列의 資料를 명확히 推定하기 위하여 定常 時系列로 변환하여 推定하는 방식이 일반적인 접근방법이다. 非定常 時系列 資料를 定常 時系列로 변환하는 방법으로는 時系列 資料의 log 또는 제곱근을 값을 취하거나 동일 資料를 일정한 지연시간을 두고 감하는 差分方式 등을 이용한다(김은영과 최병선, 1990). 本 研究에서는 給水量을 差分方式을 이용하여 時系列을 定常化 시켜 一日 給水量을 豫測하도록 하였다.

非定常 時系列로 분류되는 上水道 一日 給水量 資料를 定常 時系列로 變換하기 위하여 주어진 給水量 時系列  $\{Z_t\}$ 을 연속적으로 差分(difference)하여 평균이 일정한 定常 時系列을 얻는 방법을 택하였다. 즉,  $Z_t$ 에 대한 1차 差分은

$$W_t = (1-B)Z_t = Z_t - Z_{t-1} \quad (2)$$

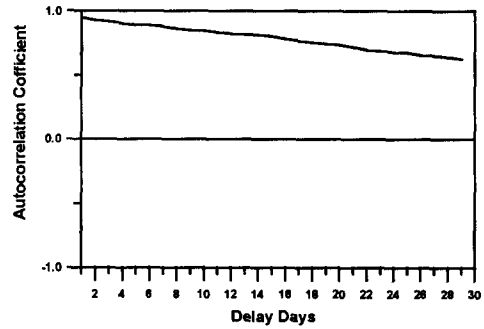


Fig. 2. Autocorrelation Coefficient of Daily Water Demand in Kwangju for the Year of 1994

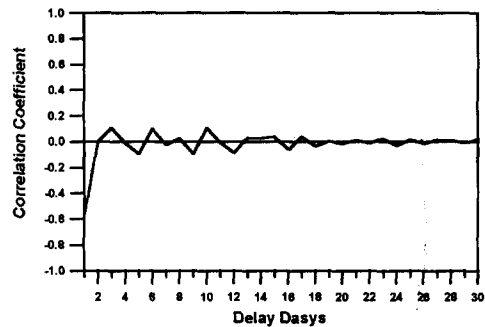


Fig. 3. Autocorrelation Coefficient of Differenced Daily Water Demand with Second Degree Lag Time

를 의미하며, 2차 差分은

$$\begin{aligned} W_t &= (1-B)^2 Z_t \\ &= (1-2B+B^2)Z_t \\ &= Z_t - 2Z_{t-1} + Z_{t-2} \end{aligned} \quad (3)$$

이다. 일반적으로 추세를 갖고 있는 非定常 時系列  $Z_t$ 에 대한  $d$ 차 差分法은  $(1-B)^d Z_t$ 로서 이를 통해 定常 時系列을 얻을 수 있다. 여기서  $d = 1, 2, \dots$  값을 갖는 差分次數이다.

差分次數를 1차에서 3차까지를 차례대로 差分하여, 자기상관계수의 수렴속도가 30일 이상 지나 수렴

하거나 수렴했다가 다시 발산하는 경우는 제외하고 수렴속도가 비교적 빠른(15일 정도) 차분차수를 선택하였다. Fig. 3은 1995년 光州廣域市 一日 給水量을 2次 差分한 時系列의 自己相關係數를 나타낸 것이다. Fig. 3의 自己相關係數를 보면 自己相關係數의 절대값이 日差가 커짐에 따라 거의 零에 수렴하고 있다. 이는 上水道 一日 給水量이 2次 差分法에 의하여 定常的인 時系列로 변환되었음을 알려주고 있다.

#### 4. 給水量 豫測 模型

##### 4.1 ARIMA 模型

상수 사용량의 예측기법으로서 이용된 ARIMA (Auto Regression Integrated Moving Average) 모형에 대해 간단히 서술하면, ARIMA 모델은 과거의 時系列 資料에 포함된 각종특성(지속성, 주기성 등)을 표현할 수 있는 모형이다. 여기에서는 과거의 給水量에 의해 현재의 상태를 선형화의 형태로 나타내는 自己回歸(AR), 과거의 백색잡음 과정에 의한 현재의 상태를 나타내는 移動平均(MA). 그리고 이것들 2개의 모델을 포함하는 自己回歸 - 移動平均(ARMA) 모형이 있다. 다시 말하면, ARMA 모형은 과거의 상태량(사용수량)에 대한 영향을 自己回歸項에 의해 과거의 時系列 변동의 평활화를 移動平均의 향으로 통합적으로 고려할 수 있는 모형이다. 또한 時系列이 定常性을 갖지 않는 資料는 差分을 해서 時系列 資料를 定常化시켜 ARMA모형을 적용하는 모델을 ARIMA모형이라고 한다(김연형, 1993 ; 최기현과 이종협, 1994).

##### 4.1.1 自己回歸(AR) 過程

自己回歸 模型은 時系列 자체에 대한 회귀형태를 취하는 모형으로 일반  $p$ 차 AR과정을 따르는  $\{Z_t\}$ 는 다음과 같이 나타낸다.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (4)$$

이때는  $a_t$ 는 평균이 0, 분산  $\sigma_a^2$ 인 백색잡음 과정이다. 따라서 현시점  $t$ 에서의 時系列  $Z_t$ 는  $p$ 개의 과거값들의 가중합과 이들의 설명되지 않는 부분인 誤差項  $a_t$ ,

의 선형결합으로 표현된다. 여기에서  $a_t$ 와  $Z_{t-1}, Z_{t-2}, \dots$ 와는 독립이라고 가정한다. 식 (4)는 후항연산자  $B$ 를 이용하여 다음과 같이 표현할 수도 있다. 즉,

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = \phi(B) Z_t = a_t \quad (5)$$

##### 4.1.2 移動平均(MA) 過程

移動平均 模型은 時系列  $\{Z_t\}$ 가 時系列 자체에 대한 회귀형태를 띠고 있는 自己回歸 과정과는 달리 현재와 과거의 백색잡음들의 가중선형결합으로 표현되는 모형으로 일반  $q$ 차 MA過程은 다음과 같이 나타낼 수 있다.

$$\begin{aligned} Z_t &= a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \\ &= (1 - \theta_1 B - \dots - \theta_q B^q) a_t \end{aligned} \quad (6)$$

이때  $\{a_t\}$ 는 평균이 0, 분산  $\sigma_a^2$ 인 백색잡음 과정이고 時系列 資料는 항상 定常性을 가지고 있어야 한다.

##### 4.1.3 自己回歸 移動平均(ARMA) 過程

時系列  $\{Z_t\}$ 가 自己回歸(AR)형태와 移動平均(MA)형태를 동시에 가지고 있다면 보다 일반적인 時系列 模型이  $\{Z_t\}$ 를 설명하는 데 필요하다. 다음과 같은 식을 만족하는  $\{Z_t\}$ 를 自己回歸 次數가  $p$ 이고 移動平均 次數가  $q$ 인 ARMA( $p, q$ ) 過程이라 정의한다.

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (7)$$

또는

$$(1 - \phi_1 B - \dots - \phi_p B^p) Z_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (8)$$

와 같이 정의된다. 여기서  $\{a_t\}$ 는 평균이 0, 분산  $\sigma_a^2$ 인 백색잡음 과정이다.

##### 4.1.4 ARIMA 模型

時系列  $\{Z_t\}$ 의  $d$ 차 差分한 時系列  $\{W_t = (1-B)^d Z_t\}$ , 또는 변수변환과  $d$ 차 差分한 時系列  $\{W_t = (1-B)^d Z_t\}$ ,

In  $Z_t$ 이 AR次數가 p, MA次數가 q인 ARMA(p, q) 模型을 갖는다면 時系列  $\{Z_t\}$ 는 次數가 (p, d, q)인 ARIMA 模型을 갖는다고 한다.

ARIMA(p, d, q) 過程은 다음과 같이 정의한다.

$$(1 - \phi_1 B - \dots - \phi_p B^p) W_t = \theta_0 + (\theta_1 B - \dots - \theta_q B^q) a_t \quad (9)$$

여기서  $W_t$ 는 적절한 변환을(差分 등) 통해 얻어진 定常時系列이고  $\{a_t\}$ 는 평균이 0, 분산이  $\sigma_a^2$ 인 백색 잡음 過程이다. 식(7)에서  $\theta_0$ 는  $d=0$ 와  $d > 0$ 일 때 서로 다른 의미를 갖는다.  $d = 0$  일때  $W_t = Z_t - \mu$ 로서 定常時系列이므로 식(7)를 후행연산자  $B$ 를 사용하면,

$$\theta(B)(Z_t - \mu) = \theta(B)a_t \quad (10)$$

인데, 이것은

$$\theta(B)Z_t = \theta_0 + \theta(B)a_t \quad (11)$$

로 표현할 수 있다. 그러므로

$$\theta_0(1 - \theta_1 B - \dots - \theta_q B^q) \mu = (1 - \theta_1 - \dots - \theta_q) \mu \quad (12)$$

로서  $\{Z_t\}$ 의 평균과 관련이 있다.

ARIMA(p, d, q) 過程은 일반적으로 식(9)과 같이 표현되며, 原時系列  $\{Z_t\}$ 를  $d$ 차差分한 時系列  $\{W_t = (1-B)^d Z_t\}$ 가 ARMA(p, q)의 過程을 따른다. 따라서 ARMA(p, q)의 過程을 따르면 된다.

## 4.2 ARIMA 模型의 構成

일반적으로 模型의 次數決定은 自己相關函數(autocorrelation function: ACF)와 偏自己相關函數(partial autocorrelation function: PACF)에 의해 결정하지만, 이렇게 하면 개인적인 주관성이 개입될 수도 있을 뿐만 아니라 自己相關函數와 偏自己相關函數로 次數를 결정하기가 애매한 경우도 있을 수 있다. 따라서 本 研究에서는 次數(p, q)를 3차까

지 조합하여 模型을 구성하여, 誤差가 많이 발생하거나 수렴이 잘되지 않은 模型은 제외시켜 模型을 구성하여 一日 上水給水量을 豫測하였다.

模型 구성을 보면 模型 1은 AMIMA(1, 2, 2), 模型 2는 ARIMA(1, 2, 3), 模型 3은 ARIMA(1, 2, 1), 模型 4는 ARIMA(2, 2, 2)차로 각각 구성하였으며, Table 2는 模型구성도를 나타내고 있다.

Table 2. Composition of Parameters for the ARIMA Model

Model Name	No.1	No.2	No.3	No.4
Autoregression(p)	1	1	1	2
Differencing(d)	2	2	2	2
Moving average(q)	2	3	1	2

## 4.3 模型의 檢證 方法

模型의 檢證 方法으로는 Portmanteau 檢證方法 및 殘差分析에 의한 方法 등이 있으나(Young, 1973; 比企三藏, 1990) 本 研究에서는 식 (13)으로 殘差分析에 의해 模型을 檢證하였다.

$$E = Z(t) - \hat{Z}(t) \quad (13)$$

여기서,  $E$ 는 殘差,  $\hat{Z}(t)$ 는 豫測值,  $Z(t)$ 는 觀測值이다.

분석은 선택한 ARIMA모형이 타당한 것이라면 殘差들은 평균이 0이고 분산이 분산  $\sigma_a^2$ 인 백색잡음과 경과 비슷한 성질을 가져야 한다. 이러한 기준을 가지고 殘差들을 분석함으로써 선택된 ARIMA모형을 검증하였다.

檢證의 첫단계는 殘差들의 時系列 산포도를 그리는 것이다. 선택된 ARIMA 모형이 타당한 것이라면 殘差들은 어떠한 특별한 형태를 보이지 않고 Y축 값이 0인 수평축 주위에 산포될 것이다. 두 번째 단계는 殘差들의 定規性을 조사하기 위해 히스토그램과 정규확률지에 그림을 그려 殘差와 정규스코어의 상관 계수를 계산한다. 정규확률지에 그림을 그리는 방법은 X축에 殘差들의 크기순으로 나열하고, t번째로 작은 殘差들에 해당되는 累積確率을 Y축에 표시한다. 이 累積確率曲線이 직선에 가까우면 誤差項은 定

規分布를 한다고 할 수 있다.

이 산포도가 직선인지 아닌지를 객관적으로 조사하는 방법중의 하나는 원래의 자료와 정규스코어의 상관계수를 사용하여 검정하는 것이다. 본 연구에서는 정규스코어들 사이의 相關係數의 채택역은 0.05의 有意水準으로 보고 모델을 檢證하였다.

### 5. 結果 및 考察

상수도 一日 給水量 推定을 위한 적절한 時系列長을 선택하기 위하여 15일, 30일, 60일, 90일 간의 給水量資料를 이용하여 正規分布에 따른 각 모델의 自己相關係數는 Table 3과 같다.

Table 3를 보면 資料의 수가 30인 경우 2월의 모델 4을 제외한 모든 모델이 0.05%의 有意水準에 정규성이 인정되며, 資料의 수가 60인 경우에도 모델 4을 제외하고 모든 모델이 正規性을 띄고 있어 棄却할 수 없는 모델로 나타났다. 그리고 資料의 수가 90인 경우에는 모든 모델에서 有意水準에 미치지 못하고 있는 것으로 나타나고 있다. 따라서 一日 給水量 豫測에 사용될 수 있는 資料의 수는 30~60개로 나타났으며, 모든 月別 資料에 대하여 有意水準을 만족하는 모델은 1, 2, 3으로 나타났다. 따라서 給水量의 추정은 추정하고자 하는 날 이전의 연속적인 30~60일간 資料를 ARIMA 模型에 이용하여 給水量을 推定하였다. 時系列長이 30일인 경우의 예를 들어보면, 4월 1일의 給水量을 추정하기 위하여 4월 1일 이전의 30일간인 3월 2일부터 3월 31일까지의 給水量 資料를 이용하였고, 다음으로 3월 3일부터 4월 1일까지의 給水量 資料를 가지고 4월 2일의 給水量을 예측하였다. 이러한 방법으로 각 모델에 대하여 一日 給水量을 豫測하였다.

給水量 推定에 사용한 모형의 信賴性은 모형의 推定結果를 이용하여 檢증하였다. 檢증방법으로는 豫測值와 實測值와의 차이 殘差를 구하고 殘差의 正規性을 檢證하는 방식으로서, 殘差가 有意水準 0.05%의 正規分布를 따르지 않은 모형의 豫測值는 棄却시키고 有意水準內에 있는 모형들의 豫測值만을 채택하였다.

Fig. 4-7은 1994년도 光州廣域市 一日 給水量에 대한 豫測值를 보여주고 있다. Fig. 4는 1994년 2월

Table 3. Correlation Coefficient between the Normal Distribuion Function and the Residual Resulting from Differences between the Estimation and the Actual Water Demand

Model Name	No.1	No.2	No.3	No.4	Critical Level $\alpha = 0.05$	Remarks
February	0.920	0.921	0.918	0.860	0.938	n=15
April	0.930	0.927	0.921	0.932		
July	0.926	0.934	0.931	0.930		
October	0.918	0.916	0.922	0.921		
February	0.967	0.970	0.974	0.917	0.964	n=30
April	0.989	0.981	0.982	0.995		
July	0.989	0.992	0.989	0.988		
October	0.980	0.960	0.982	0.989		
February	0.976	0.975	0.982	0.923	0.976	n=60
April	0.980	0.983	0.981	0.975		
July	0.995	0.990	0.995	0.971		
October	0.976	0.989	0.991	0.971		
February	0.967	0.968	0.970	0.927	0.985	n=90
April	0.982	0.982	0.978	0.975		
July	0.984	0.985	0.986	0.980		
October	0.982	0.989	0.987	0.982		

\* n = 豫測에 사용된 資料의 수(時系列長)

에 대한 給水量 추정 결과를 보여주고 있다. 2월의 給水量 추정에 관한 모형의 信賴性을 檢증한 결과 0.05%의 유의 수준에서 모형 4를 제외한 모형 1, 모형 2, 모형 3은 信賴性이 檢증되었다. 모형 1, 모형 2와 모형 3은 변화가 많지 않은 날들의 給水量 資料에 대하여 비슷한 정도의 推定을 하는 것을 확인할 수 있으나 給水量의 변화가 심하였던 2월 12일에 실제 급수량과 예측치 간의 차이가 심하게 보이고 있으며 豫測值와 實測值와 간의 推定誤差는 3만톤 정도에 이르고 있다. 一日 給水量 32만톤에 비하여 3만톤의 誤差는 10%미만의 誤差로서 상수도시설 운영상 용인할 수 있는 誤差率이라고 하나 給水量이 급변하는 부분에서의 ARIMA모형에 의한 推定은 어려움이 있는 점을 시사하고 있다.

Fig. 5는 봄철인 4월의 1일 給水量에 대한 實測值와 豫測值이다. 4월의 경우는 2월의 경우와 달리 給水量의 변화의 폭이 작은 관계로 모형에 의한 豫測值는 실제급수량과 매우 근접함을 보여 주고 있다. 35만톤 내외의 給水量에 대한 豫測誤差도 1.5만톤 내외

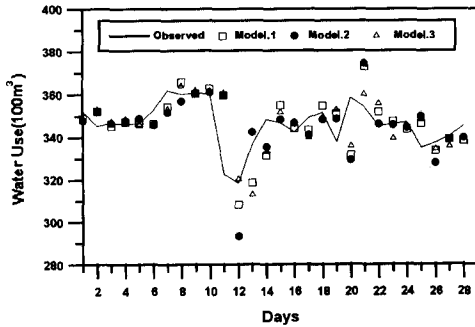


Fig. 4. Estimation of Daily Urban Water Demand Using ARIMA Model for February, 1994

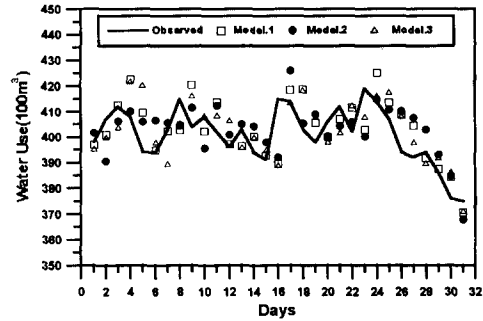


Fig. 7. Estimation of Daily Urban Water Demand Using ARIMA Model for October, 1994

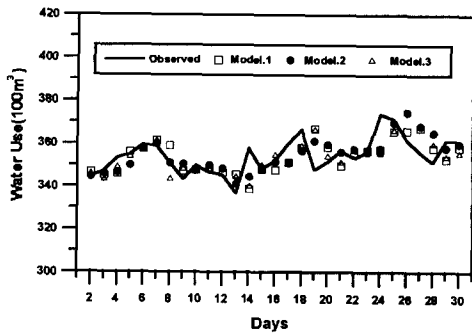


Fig. 5. Estimation of Daily Urban Water Demand Using ARIMA Model for April, 1994

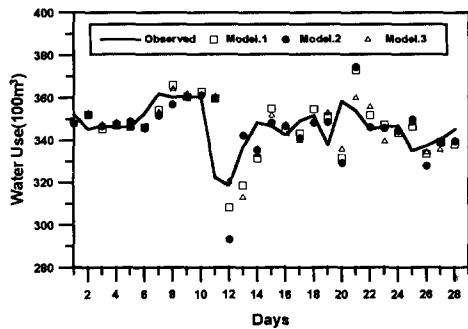


Fig. 6. Estimation of Daily Urban Water Demand Using ARIMA Model for July, 1994.

로서 만족할 만한 推定結果를 보이고 있다.

Fig. 6은 여름철인 7월에 대한 實測値와 豫測値의 결과이다. 기온 상승으로 인한 給水量 증가로 인하여 7월의 給水量은 一日 46만톤의 규모로 상승하고 있다. 7월중 상수도시설 운영상 사고 등으로 인한 단수 조치가 없었던 관계로 給水量 변화의 폭이 작은 것으로 나타났으며, 7월의 給水量에 대한 豫測値도 4월과 마찬가지로 實測値와 잘 맞고 있음을 알 수 있다.

Fig. 7은 가을철인 10월에 대한 觀測値와 豫測値의 결과이다. 10월도 4월 및 7월과 마찬가지로 豫測値이 觀測値와 비교적 잘 일치하고 있음을 보여주고 있다.

Table 4는 상수도 一日 給水量의 豫測値에 대한 誤差率을 정리하였다. 誤差率은 모형의 信賴性이 확보된 모형 1, 2 와 3에 대하여 時系列長(資料의 수)이 30과 60인 경우를 정리하였다. 모형의 정도를 비교하기 위하여 誤差率이 크게 발생한 2월에 대하여 時系列長이 30인 경우를 선택하여 모형별 誤差率을 비교하면 모형 3에서 誤差率은 11.7% 으로서 最大誤差率을 기록하고 있으며, 모형 1과 2도 모형 3의 경우보다 0.3% 내지 0.4% 정도 작으나 이와 비슷한 誤差率을 보여주고 있다. 2월중 給水量에 대한 각 모형의 平均誤差率을 보면 약 2.4%로서 3가지 모형 모두 비슷한 정도를 갖는 모형임을 보여주고 있다.

또한 모형이 어느정도까지 정확할 수 있는지 알기 위하여 最小誤差率을 조사한 결과 2월달에 대한 각 모형의 誤差率은 0.2% 미만으로 나타나고 있다.



Table 4. Errors of ARIMA Model Estimating the Daily Urban Water Demand for the Year of 1994

Month	Model No.1			Model No.2			Model No.3			Remarks
	Max	Min	Ave	Max	Min	Ave	Max	Min	Ave	
February	11.4593	0.0141	2.3990	11.4250	0.0084	2.4013	11.7121	0.1278	2.3881	n = 30
April	5.9026	0.0606	1.8546	4.2753	0.0384	1.4825	5.3284	0.2965	1.7311	
July	5.1774	0.0209	1.8975	5.9122	0.2409	2.0782	5.1771	0.3709	0.2887	
October	6.1638	0.1602	2.0571	5.5313	0.1604	2.1744	6.6228	0.0073	2.2324	
February	11.9791	0.0121	2.5671	11.9214	0.0101	2.5313	12.0145	0.1415	2.4128	n = 60
April	6.1832	0.0385	1.6797	6.6685	0.0131	1.8256	6.6137	0.2965	1.9291	
July	6.0454	0.1017	1.8568	6.2626	0.2036	1.9851	6.1248	0.0095	2.0576	
October	6.9125	0.2119	2.0230	6.1267	0.1824	2.5781	6.8924	0.0149	2.6714	

\* n = 豫測에 사용된 資料의 수(時系列長의 길이)

4월의 給水量 추정에 관한 모형의 誤差率을 時系列長이 30인 경우에 대하여 조사한 결과 最大誤差率은 모형 1에서 약 6%가 발생하였으며 모형 2와 3에서 약 4 ~ 5%의 誤差率을 보이고 있다. 平均誤差率 및 最小誤差率은 2월과 비슷하게 나타났다. 7월의 給水量 추정에 대한 最大誤差率은 모형 2에서 약 6%로 나타났으며, 나머지 모델들에서는 약 5%의 誤差率을 보이고 있다. 平均誤差率 및 最小誤差率은 2월 및 4월과 거의 비슷하게 나타났다. 10월의 最大誤差率은 모형 3에서 약 6.6%로 最大值이며, 나머지 모델들에서는 약 6%의 誤差率을 보이고 있다. 平均誤差率 및 最小誤差率은 2월, 4월 및 7월과 비슷하게 0.2% 미만으로 나타났다.

豫測에 사용된 資料의 수인 時系列長이 60인 경우를 보면, 最大誤差率面에서는 豫測資料의 수가 30인 경우보다도 약 1~2% 정도의 誤差가 더 발생하는 것으로 나타난다. 이는 豫測日 給水量은 예측일에서 멀어질수록 영향을 적게 받고 있는데 너무 時系列長이 길면 自己回歸값이 분산되기 때문인 것으로 판단된다.

最小 및 平均誤差率은 豫測에 사용된 資料의 길이에 상관없이 거의 일정함을 알 수 있다. 따라서 一日 上水給水量 豫測에 사용할 수 있는 資料의 수는 약 30개 정도가 最高의 程度을 얻을 수 있을 것으로 판단되었다.

하지만 본 논문은 時系列長을 15, 30, 60, 90일에 대해 각각 모델을 적용하여 적정 時系列長을 결정하였지만 다른 도시에 ARIMA 모형을 적용하려면 자

기상관계수에 의한 검토를 하고 이어 자기상관계수에 의한 통계적 유의성을 검증하여 時系列長을 결정하는 것이 시행착오로 時系列長을 결정하는 것보다 시간이 절약될 것으로 판단된다.

## 6. 結 論

本 研究에서는 上水道 一日 給水量 推定을 위하여 ARIMA모형을 구성하고, 대상지역을 光州廣域市를 선정하여 조사한 결과 다음과 같은 結論을 얻었다.

(1) 상수도 一日 給水量豫測을 위한 ARIMA 모형의 構成을 위한 時系列資料의 定常化는 2차 差分을 함으로써 얻을 수 있었다.

(2) 상수도 一日 給水量豫測을 위한 ARIMA 모형에서 時系列長은 30일이 모형의 精度를 가장 높일 수 있었다.

(3) ARIMA 모형 구성은 自己回歸( 전달함수)項의 次數는 1차에서 정도가 높고, 自己回歸( 전달함수)項의 1차에 대한 移動平均項의 次數는 1차에서 3차까지 次數變化를 하여도 一日 給水量을 豫測하는데 큰 영향을 미치지 않는 것으로 나타났다.

(4) 本 研究에서 이용한 ARIMA 모형은 上水道 一日 給水量에 영향을 미치는 最高氣溫, 濕度, 降雨量 등의 氣候因子를 고려하지 않고도 최대 12%, 평균 2%의 오차를 얻어 수용할 만하나, 앞으로 이들 인자의 영향을 고려할 수 있는 多變量 ARIMA모형에 대한 研究가 필요하다고 생각된다.

## 참 고 문 헌

- 김수원 (1985). “都市用水 計劃을 위한 물 需要 豫測에 관한 研究,” 박사학위 논문, 영남대학교.
- 김연형 (1993). 時系列分析과 豫測. 自由아카데미.
- 김은영, 최병선 (1990). 미니팁을 이용한 時系列 분석입문.
- 임동조 (1995). “氣候變動에 따른 上水道 一日 給水量 豫測에 관한 研究,” 석사학위 논문, 전남대학교.
- 이경훈, 박성천, 문병석 (1994a). “上水道의 給水量 및 配水池 容量變化에 관한 研究.” 대한상하수도학회, 제9권, 제2호, pp. 118-126.
- 이경훈, 이삼노, 문병석 (1994b) “上水道 1日 給水量 變化와 確率分散 模型에 관한 研究.” 제36회 수공학 연구발표회 논문집, 한국수문학회, pp. 405-410.
- 이경훈, 이삼노, 문병석 (1994c). “上水道 1日 給水量的 時間的 變化의 特性에 관한 研究.” 한국수문학회지, 제27권, 제2호, pp. 135-143.
- 최기현, 이종협 (1994). SAS/ETS를 이용한 時系列 分析과 그 應用. 自由아카데미.
- 현인환, 목동우 (1995). “상수도 급수량의 단기예측.” 대한상하수도학회지, 제9권, 제1호, pp. 96-106.
- 比企三藏 (1990). “カルマンフィルタによる配水量豫測.” 日本水道協會雜誌, 第58卷, 第2號, pp. 2-14.
- 小泉明, 千田孝一, 川口十郎 (1988). “多元ARIMAモデルによる水使用量の短期豫測.” 日本水道協會雜誌, 第57卷, 第12號, pp. 13-20.
- Frankel, R.J., and Shouvanavirakul, P. (1973). “Water consumption in small communities of northeast Thailand.” *Water Resources Research*, Vol. 9, No. 5, pp. 1196-1207.
- Maidment, D.R., and Parzen, E. (1984). “Cascade model of monthly municipal water use.” *Water Resources Research*, Vol. 20, No. 1, pp. 15-23.
- Young, R.A. (1973). “Price elasticity of demand for municipal water: A case study of Tuscon, Arizona.” *Water Resources Research*, Vol. 9, No. 4, pp. 1068-1072.

〈접수: 1996년 4월 15일〉