

氣候條件에 의한 上水道 一日 給水量的 變化에 관한 研究 A Study on the Variation of Daily Urban Water Demand Based on the Weather Condition

李 庚 焄* · 文 炳 錫** · 嚴 東 祚***

Rhee, Kyoung Hoon · Moon, Byoung Seok · Eom Dong Jo

Abstract

The purpose of this study is to establish a method of estimating the daily urban water demand using statistical model. This method will be used for the development of the efficient management and operation of the water supply facilities. The data used were the daily urban water use, the population, the year lapse and the weather conditions such as temperature, precipitation, relative humidity, etc. Kwangju city was selected for the case study area. The raw data used in this study were rearranged either by month or by season for the purpose of analysis, and the statistical analysis was applied to the data to obtain the regression model. As a result, the multiple linear regression model was developed to estimate the daily urban water use based on the weather condition. The regression constant and the model coefficients were determined for each month of a year. The accuracy of the model was within 3% of average error and within 10% of maximum error. The developed model was found to be useful to the practical operation and management of the water supply facilities.

요 지

본 논문에서는 상수도시설을 效率的으로 운영하는 데 필요한 1日 給水量 需要를 豫測하는 방식에 대하여 統計學的 模型概念을 이용하여 연구하였다. 給水量的 變化는 氣候와 많은 연관성이 있는 것으로 알려지고 있으나 給水量的 豫測과 상수도시설의 운영에 있어서 給水量 變動要因을 고려한 數學的 혹은 統計學的 모형을 개발하여 실용화하지 못하고 있다. 따라서 본 연구에서는 시간이 지남에 따라 給水量이 증가하는 경향을 반영할 수 있는 經年變化, 給水人口와 더불어 上水道 給水量的 중요한 變動要因으로 간주되는 氣候因子를 이용하여 上水道 1日 給水量을 예측할 수 있는 모형을 제안하고 그 有效性을 檢證 하였다. 給水量 예측모

* 전남대학교 토목공학과 부교수

** 전남대학교 토목공학과 박사과정 수료

*** 전남대학교 토목공학과 석사과정

형식은 統計學的 方法중에서 回歸模型式을 사용하였으며, 연구대상 지역으로서는 자료수집이 용이한 光州廣域市를 선정하였다. 1日 給水量을 예측하는데 있어서 回歸模型의 시계열구간을 年別, 月別, 季節別로 분류하고 시계열 구간별로 최적의 回歸模型式을 제안하고자 回歸分析을 시행하였다. 回歸分析을 시행하기 위해서 먼저 給水量을 從屬變數로 하고 給水人口, 經年變化 및 氣候因子 즉, 최고기온, 평균기온, 강우량, 운량, 일사량, 상대습도, 일조시간 등을 說明變數로 하는 相關分析을 행하고 비교적 相關性이 높은 인자를 택하여 모형식을 구성하였다. 제안된 모형식은 사고 등에 의한 단수조치나 한발의 영향으로 인한 隔日制給水를 실시하는 시기를 제외하고 實測値와 模型과의 誤差率이 최대 10%, 평균 3% 이내로 나타났으며, 모형식의 결과는 上水道 1日 給水量 예측이 필요한 시설에 적용 가능하다고 판단되었다.

1. 序 論

産業構造가 기반산업 위주에서 고도의 工業爲主로 전환되고, 급격한 經濟成長을 이루면서 국민의 생활수준이 크게 향상되고, 人口가 都市에 편중되어 도시의 비대화를 초래하고 있다. 이러한 현대사회는 복합적인 問題點을 유발하고 있으며 물 需要量도 생활용수, 영업용수, 공업용수, 농수산업용수 등 다방면에 걸쳐서 需要量이 증가하여, 기존의 用水로는 수량면에서 수요량을 충족시키지 못하고 있다.

따라서 상수도의 과학적이고 합리적인 운영을 통해서 水量 및 水質面에서 수요자의 요구에 맞는 질 높은 수도 서비스를 제공하기 위하여 需要量을 예측하고 이에 따라서 상수도시설을 유효·적절하게 운용하는 것이 중요하다.

上水道 1日 給水量을 예측하는 방식은 과거의 자료로부터 統計學的 方法에 의하여 예측할 수 있는 回歸分析法과 時系列分析法으로 접근할 수 있다. 또한 給水量 예측은 시점에 따라 短期豫測과 長期豫測으로 분류할 수 있다. 短期豫測은 예측기간을 시간단위나 日단위로 구분하여 給水量을 추정하는 것으로 추정된 給水量은 급·배수량을 실시간 제어하는 데 유효하며, 長期豫測은 예측구간을 年別, 季節別 또는 月別로 구분하여 平均給水量을 예측하고, 예측된 자료는 급·배수시설의 효율적인 운영계획에 필수적인 자료가 된다.

回歸分析法이나 時系列分析法, 또는 短期豫測이나 長期豫測에 있어서 氣候因子는 給水量의 變化에

영향을 주는 중요한 因子로 인식되고 있으며, 이는 給水量과 氣候因子와의 相關關係를 규명함에 있어 기온과 강우량 등의 氣候因子를 도입한 給水量 연구가 처음 시도될 때 확인되었다(Howe와 Linaweaver, 1967). 그 이후 年別로 변화되는 給水量의 자료를 해당도시의 인구수, 세대의 평균수입, 수도요금과 더불어 강우량, 기온, 증발량과 연관시키려는 연구를 시도하였으며(Young, 1973), 季節別 급수수요모형을 1차선형 多重回歸分析방법으로 계산하였다(Frankel과 Shouvanavirakul, 1973). 또한 月別 平均給水量을 강우량, 증발량, 기온 등 氣候因子를 사용하여 1日 給水量을 예측한 바 있다(Maidment와 Parzen, 1984).

國內의 연구동향을 보면, 都市給水에 영향을 미치는 요인을 說明變數로 하여 用途別로 需要豫測을 하였고(김수원, 1985), 上水道 1日 給水量의 時間別, 曜日別, 季節의 분포에 관한 연구를 하였으며(이경훈 등, 1994a; b), 인구수, 최고기온, 상대습도, 강우량 등을 이용한 上水 給水量의 短期豫測을 시도한 바 있으나(현인환과 목동우, 1995), 給水量豫測과 이의 활용에 대한 國內의 연구는 아직 미흡하여 앞으로 이 분야의 연구가 더욱 활발히 추진되어야 할 것으로 사료된다. 따라서 本 研究에서는 光州廣域市의 1日 上水 需要量의 추정을 위해 給水量에 영향을 주는 變動要因으로 給水人口, 經年變化 및 氣候因子를 사용한 1日 給水量 豫測 模型式의 적용성을 검증하여 상수도시설을 합리적이고 과학적으로 운용하는 데 필요한 자료확보 및 운영방법의 개선을 도모하고자 한다.

2. 給水量 豫測理論

2.1 多重回歸分析

回歸分析法(regression analysis)은 從屬變數와 說明變數간의 함수적인 關係를 규명하기 위해서 임의의 數學的 模型을 가정하고, 이 模型을 측정된 變數들의 자료로부터 실측된 자료에 적합한 회귀방정식에 접근시켜 회귀계수를 추정하는 統計分析의 한 방법이다.

從屬變數가 하나이며 여러개의 說明變數가 있을 때 說明變數들이 從屬變數에 어떻게 영향을 미치는가를 분석하는 것이 多重回歸分析(multiple regression analysis)이다.

k 개의 說明變數 X 와 연관지을 수 있는 從屬變數 Y 가 N 개 존재하는 모집단을 생각해 보면 다음과 같이 나타낼 수 있다.

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \dots + \beta_k X_{1k} \\ Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \dots + \beta_k X_{2k} \\ &\vdots \\ Y_N &= \beta_0 + \beta_1 X_{N1} + \beta_2 X_{N2} + \dots + \beta_k X_{Nk} \end{aligned} \quad (1)$$

여기서 β_0 는 회귀상수, $\beta_1, \beta_2, \dots, \beta_k$ 는 회귀계수이며 $\epsilon_1, \epsilon_2, \dots, \epsilon_N$ 는 誤差항이다.

식 (1)을 행렬을 사용하여 표기하면,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{N1} & X_{N2} & \dots & X_{Nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix} \quad (2)$$

이다. 따라서

$$\underline{Y} = \underline{X} \underline{\beta} + \underline{\epsilon} \quad (3)$$

$(N \times 1) \quad \{N \times (k+1)\} \quad \{(k+1) \times 1\} \quad (N \times 1)$

오차항 조건은, $E(\epsilon_i) = 0$, $Var(\epsilon_i) = 0$, 그리고 $Cov(\epsilon_i, \epsilon_j) = 0$ 으로서 독립적인 정규확률변수이다. $E(\underline{Y}) = \underline{X} \underline{\beta}$ 이며 표본자료에서 구할 수 있는 \underline{b} 를 $\underline{\beta}$ 의 추정벡터라 하면,

$$\underline{\hat{Y}} = \underline{X} \underline{\beta} \quad (4a)$$

$$\underline{Y} = \underline{X} \underline{b} + \underline{e} \quad (4b)$$

이며 \underline{e} 는 實測値와 推定値의 차이인 잔차벡터이다.

모집단에서의 回歸係數 벡터의 $\underline{\beta}$ 와 誤差分散 σ^2 의 값은 n 개의 표본자료를 이용하여 잔차자승의 합을 최소로 하는 \underline{b} 로부터 추정할 수 있다.

$$\begin{aligned} \sum_{i=1}^n e_i^2 = \underline{e}' \underline{e} &= (\underline{Y} - \underline{X} \underline{b})' (\underline{Y} - \underline{X} \underline{b}) \\ &= \underline{Y}' \underline{Y} - 2 \underline{b}' \underline{X}' \underline{Y} + \underline{b}' \underline{X}' \underline{X} \underline{b} \\ &= \underline{Y}' \underline{Y} - \underline{Y}' \underline{X} \underline{b} \end{aligned} \quad (5)$$

이다. 이것을 \underline{b} 에 관하여 미분한 후에 0으로 놓으면,

$$\frac{\partial \underline{e}' \underline{e}}{\partial \underline{b}} = -2 \underline{X}' \underline{Y} + 2 \underline{X}' \underline{X} \underline{b} = 0$$

이다. 따라서 多重回歸模型의 계수벡터 \underline{b} 에 관하여 정리하면 다음과 같다.

$$\underline{b} = (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{Y} \quad (6)$$

2.2 回歸模型의 適合性

統計學的인 適合性檢定 방식은 계산된 回歸方程式이 統計的으로 有意한가를 검정하는 것이다. SST (total sum of squares)는 SSE(sum of squares due to errors)와 SSR(sum of squares due to regression)로 분류되며, 適合性檢定을 위한 分散分析(Analysis of variance)은 SSE 變動平均(MSR)과 SSR 變動平均(MSE)의 비율을 계

산한 F 값을 臨界値와 비교하여 결정한다. 자료의 SST을 계산하면,

$$\begin{aligned} SST &= \sum(Y_i - \bar{Y})^2 = \sum Y_i^2 - n(\bar{Y})^2 \\ &= \underline{Y}' \underline{Y} - n(\bar{Y})^2 \end{aligned} \quad (7)$$

이다. 다음으로 SSR을 구하면,

$$\begin{aligned} SSR &= \sum(\hat{Y}_i - \bar{Y})^2 = \sum \hat{Y}_i^2 - n(\bar{Y})^2 \\ &= \underline{\hat{Y}}' \underline{\hat{Y}} - n(\bar{Y})^2 \end{aligned} \quad (8a)$$

이다. 여기에 식 (4)와 (6)을 대입하여 정리하면,

$$\begin{aligned} SSR &= \{(X'X)^{-1}XY\}' X'X \{(X'X)^{-1}XY\} \\ &\quad - n(\bar{Y})^2 = b' X' Y - n(\bar{Y})^2 \end{aligned} \quad (8b)$$

이며

$$\begin{aligned} SSE &= \sum(Y_i - \hat{Y}_i)^2 = \sum(Y_i - \hat{Y}_i)(Y_i - \hat{Y}_i) \\ &= \underline{Y}' \underline{Y} - b' X' Y \end{aligned} \quad (9)$$

이다. 自由度를 계산하면, SST는 일반적으로 $n-1$ 개의 自由度를 갖는다. SSE는 $n-(k+1)$ 의 自由度를 가지며, SSR은 獨立變數 k 개의 갯수만큼 自由度를 갖는다. 위에서 얻어진 자승합과 自由度를 이용하여 分散分析에 사용되는 分析表를 만들면 Table 1과 같다.

또한, 從屬變數 Y 와 k 개의 說明變數 집합 X_1, X_2, \dots, X_k 사이의 관계를 檢定해야 한다. 이를 위해서는 먼저 回歸模型에 관한 假說을 세우고 n 개의 표본자료에 대한 F 값을 계산한 후 自由度를 고려한 臨界値와 비교하여 假說의 채택 여부를 결정한다.

有意水準 α 에서, $F \leq F_{(\alpha, k, n-k-1)}$ 이면 귀무가설 (Null hypothesis)이 採擇되고, $F > F_{(\alpha, k, n-k-1)}$ 이면 귀무가설 (Null hypothesis)이 棄却된다.

2.5 回歸模型의 精度 및 相關度

回歸模型이 어느 정도로 실측된 자료를 잘 설명하고 있는가 하는 문제는 回歸模型의 多重決定係數

(R^2 : coefficient of multiple determination)를 살펴봄으로써 알 수 있다. 그런데 R^2 값에 대한 해석은 주의할 필요가 있다. R^2 이 크다고 해서 반드시 적합한 模型이라고는 볼 수 없는데, 이는 어떤 豫測의 경우, R^2 이 크다 하더라도, MSE가 너무 커서 추론이 불 가능할 수도 있기 때문이다. 따라서 이러한 폐단을 막기 위하여 決定計數를 수정변경하여 측정한, 修正決定係數(R_a^2 : adjusted coefficient of determination)를 써야 한다.

$$R_a^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} = 1 - \left(\frac{n-k}{n-k-1}\right) \left(\frac{SSE}{SST}\right) \quad (10)$$

R_a^2 는 說明變數의 수가 늘어난다 하더라도, 이를 보완하여 決定係數값에 영향을 주지 않는 장점을 갖고 있어 模型의 적합성을 판단하는 중요한 因子로 인식되고 있다.

다음으로 從屬變數와 說明變數, 說明變數간의 相關性을 검토해야 한다. 주관심이 되는 두 變數의 선형관계의 방향과 정도는 相關係數(correlation coefficient)를 이용하여 측정하는데 相關係數 ρ 는 식 (11)과 같다.

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \quad -1 \leq \rho \leq 1 \quad (11)$$

여기서, σ_{xy} 는 X 와 Y 두 變數의 共分散이며 σ_x 와 σ_y 는 각각 X 와 Y 의 標準偏差이다.

3. 對象地域 및 研究資料

本 研究의 대상지역으로는 우리나라의 5대도시 중의 하나인 光州廣域市를 선정하였고, 給水量資料와 氣候資料를 확보하여 本 研究에 활용하였다. 光州廣域市의 1994년 12월말 현재 人口는 1,274,000여명에 이르고 있으며, 面積은 약 500km², 個人別 GNP가 \$6,000에 달하는 도시로서 光州廣域市에 파악된 上水道 給水現況과 조사방식은 우리나라에서 이와 비슷한 규모의 중, 대형도시에 적용할 수 있으리라 생각된다.

Table 1. Analysis of Variance for Regression Model.

Source	Sum of Squares	Degree of Freedom	Average of Sum of Squares	F
Regression	$SSR = \sum (Y - \bar{Y})^2$	k	$MSR = \frac{SSR}{k}$	$\frac{MSR}{MSE}$
Residual	$SSE = \sum (Y - \hat{Y})^2$	$n - (k + 1)$	$MSE = \frac{SSE}{n - k - 1}$	
Total	$SST = \sum (Y - \bar{Y})^2$	$n - 1$		

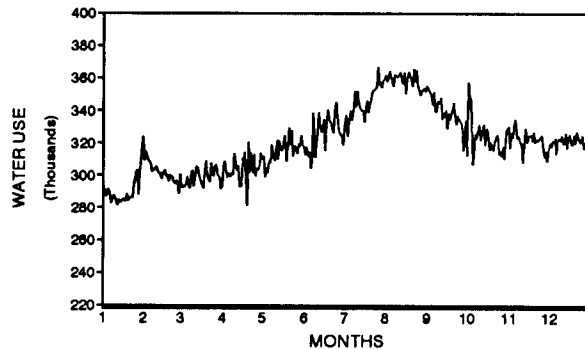


Fig. 1 The Variation of Daily Water Use Kwangju City for the Year of 1990.

光州廣域市の 給水狀況을 살펴보면 1994년 말에 平均 38만톤/일 규모의 給水量에 달하였고, 用途別 給水量을 보면 가정용 69.78%, 公業용이 4.32%, 공동목욕탕을 포함한 영業용이 21.12%, 公公 및 기타용이 4.78%이다. 光州廣域市の 산업구조상 公業용수의 給水量이 차지하는 비율이 다른 用途에 비하여 낮은 것이 특징이다. 本 研究에서 一日 給水量의 豫測에 사용한 資料는 1990년~1992년 및 1994년의 4개년에 해당하는 光州廣域市の 給水量에 관한 資料이다. 1993년의 給水量資料는 한발로 인한 水源池의 저수율 강하로 장기간 斷水措置를 실시했기 때문에 給水量의 변화가 평균값과 현저하게 차이는 관계로 해석에 사용되는 資料에서 제외하였다. Fig. 1은 光州廣域市 上水道事業本部에서 측정한 1990년 光州廣域市の 給水量을 표시하고 있다.

給水量의 추정 說明變數인 氣候因子로는 최고기온, 평균기온, 상대습도, 강우량, 증발량, 운량, 일조시간, 일사량을 사용하였다. 이들 氣候資料는 광

주지방 기상청에서 관측된 자료를 사용하였고, 給水人口는 給水量을 豫測하는데 중요한 因子이지만, 발표된 人口資料는 年末을 기준으로 하는 統計値이고, 光州廣域市와 같이 人口가 계속 증가하고 있는 都市의 경우 年末의 給水人口를 같게 보는 것은 무리가 있다고 판단되어 給水人口가 等差의으로 증가한다고 가정한 추정된 給水人口를 사용하였고, 생활수준과 문화수준의 향상에 따라 一人 一日 給水量은 매년 증가하고 있으므로 經年變化를 고려하여 給水量을 예측하였다.

4. 分析結果 및 考察

4.1 主要因子的 相關分析

本 研究에서는 上水道 給水量의 回歸模型을 개발하기 위해서 回歸模型의 구성에 필요한 說明變數로 최고기온, 평균기온, 상대습도, 강우량, 평균운량, 일조시간, 일사량 등의 氣候因子와 給水人口 및 經年變化를 선정하여 說明變數 상호간에 相關關係와

독립성을 확인하고자 相關分析을 실시하였다. 相關分析에 이용한 자료에 대해 시계열구간을 月別, 季節別, 年別區間으로 분류하여 相關係數를 구하였다.

4.1.1 氣溫

給水量과 氣候因子 중에서 氣溫과의 相關性이 가장 높게 나타났으며, 年別과 季節別, 月別區間에서 최고기온, 평균기온의 相關係數는 0.6~0.7 정도로 나타났고, 최고기온이 평균기온보다 약간 더 높게 나타났다. 또한 說明變數간의 相關係數는 최고기온과 평균기온 사이에 높은 相關性(0.8~0.92)이 존재하였다. 따라서 給水量 豫測模型에서는 최고기온을 說明變數로 채택하였다.

4.1.2 降雨量, 濕度 및 雲量

강우량과 습도, 운량은 하나의 계열로 간주하고 이들 因子가 給水量에 미치는 영향과 인자상호간의 相關性을 조사하였다. 강우량, 습도 및 운량이 給水量에 미치는 영향을 서로 비교하면 강우량의 영향은 작으며 습도와 운량은 비슷한 정도의 영향을 미치고 있다. 이들 因子와 給水量과의 相關度는 逆

相關關係이며, 給水量에 대한 이들의 영향은 미소하고, 季節적으로 다르게 영향을 미치고 있는 것으로 나타나고 있다. 봄, 가을, 겨울철에는 이들 因子와 給水量과의 相關度는 相關係數가 -0.25보다 작게 나타나고 있으나, 여름철에는 이들 因子와 給水量과의 相關度가 높아지는 것으로 나타나고 있다.

給水量에 대한 습도의 相關度를 보면 겨울, 봄철에는 給水量과의 相關度가 거의 -0.1 미만으로 작게 나타나고 있으나 6월부터 相關度가 상승하기 시작하여 8월에는 -0.4까지 逆相關度가 높아지다가 9월부터는 다시 감소하는 경향을 보이고 있다. 또한 강우량과 습도 및 운량의 相關係數는 0.6~0.8사이로 높은 相關性을 보였다.

4.1.3. 日照時間, 日射量

일조시간과 일사량은 같은 계열로 묶고 給水量과의 相關性을 비교하였다. 일조시간과 일사량은 서로 비슷하게 給水量의 變化에 영향을 주는데 일조시간이 일사량의 영향보다 약간 높게 나타났다.

일조시간과 일사량의 相關度는 相關係數가 평균 0.9정도로 두 說明變數간에 相關性은 매우 높은 것으로 조사되었다. 특히, 일조시간과 일사량은 여름

Table 2. Correlation Coefficients Obtained Using the Monthly Data for the Years of 1990 to 1994.

Months	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉
Jan.	0.870	0.113	0.001	-0.033	-0.088	0.097	0.109	0.142	0.922
Feb.	0.848	0.090	-0.252	-0.185	-0.028	0.185	0.172	0.170	0.846
March	0.811	0.221	-0.123	-0.045	0.091	0.022	0.065	0.282	0.873
April	0.835	0.320	-0.084	-0.018	-0.024	0.068	0.056	0.348	0.884
May	0.842	0.391	-0.059	-0.054	-0.111	0.271	0.153	0.483	0.885
June	0.795	0.489	-0.194	-0.140	-0.235	0.018	0.103	0.518	0.837
July	0.911	0.670	-0.383	-0.166	-0.614	0.498	0.456	0.687	0.897
Aug.	0.903	0.613	-0.298	-0.206	-0.154	0.203	0.160	0.629	0.896
Sept.	0.920	0.472	-0.197	-0.089	-0.346	0.300	0.283	0.524	0.920
Oct.	0.946	0.343	0.072	-0.043	0.136	0.088	0.129	0.445	0.941
Nov.	0.924	0.215	-0.176	-0.026	-0.032	0.004	0.002	0.324	0.919
Dec.	0.918	0.148	-0.182	-0.084	-0.005	0.027	0.028	0.169	0.913

Note that X₁: Population, X₂: Average temperature(°C), X₃: Relative humidity(%),

X₄: Precipitation(mm), X₅: Cloudiness, X₆: Sunshine duration(hr),

X₇: Solar radiation(M.j/m²), X₈: Maximum temperature(°C), X₉: Lapse of year,

β₀: Regression constant, β₁, β₂, ..., β₉: Regression coefficient

Table 3. Regression Model for Predicting the Municipal Water Use based on the Weather Condition.

	Regression Model
Basic Model	$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9$
Model No. 1	$\beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9$
Model No. 2	$\beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9$
Model No. 3	$\beta_0 + \beta_1 X_1 + \beta_4 X_4 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9$
Model No. 4	$\beta_0 + \beta_1 X_1 + \beta_4 X_4 + \beta_6 X_6 + \beta_8 X_8 + \beta_9 X_9$

Table 4. The Regression Coefficients of Model No. 4 Obtained for Each Monthly Data for the Years of 1990 to 1994.

Months	β_0	β_1	β_4	β_6	β_8	β_9	R
Jan.	351265	535210	281	-556	-265	-23571	0.933
Feb.	-2714112	502061	-183	-644	235	5356	0.909
March	7363087	1175947	118	-73	399	30239	0.913
April	7419864	1184457	46.8	101	300	30850	0.913
May	11136746	1802661	-53	-311	730	40534	0.929
June	14247981	2318064	77.8	-182	446	48258	0.895
July	-6786701	1170329	100	-334	3256	-2332	0.942
Aug.	-16465230	2797394	-0.8	278	231	-33791	0.917
Sept.	-1954871	372340	-11.6	-527	2074	11554	0.946
Oct.	-17894278	3024818	283	-726	1799	-33871	0.958
Nov.	-8439922	1454433	83	-468	1132	-11885	0.941
Dec.	-5526311	972157	-380	-237	262	-4473	0.919

철인 7, 8월에 相關性이 가장 높다. 이것은 여름철의 낮길이에 의한 영향과 태양에 의한 낮동안의 지표면 복사열에 기인하는 것으로 판단된다.

4.1.4 給水人口

光州廣域市の 人口는 매년 증가하는 추세에 있다. 이러한 증가추세에 있는 人口數에 上水道 보급율을 곱한 수를 相關分析에서 사용한 給水人口數로 사용하고 있다. 給水量과 給水人口의 相關度를 조사한 바에 의하면 두 因子의 相關係數는 0.8이상이었다. 人口數의 증감에 따라 給水量은 많은 영향을 받는 것으로 판단되었으며, 給水人口는 回歸分析에서 중요한 說明變數 역할을 할 것으로 사료된다.

4.1.5 經年變化

上水道 給水量은 해마다 증가하는 경향이며, 이러한 성향은 給水人口의 증가와 더불어 一人 1日 給水量의 使用水量에 대한 영향으로 알려져 있다.

이와 같이 都市의 발전이 거의 완속단계에 있거나 발전이 느리고 소득수준이 매우 높은 경우에는 이러한 영향이 크게 변하지 않겠지만 光州廣域市の 경우처럼 해마다 1人 1日 給水量이 증가하고 있는 성장추세는 總 平均給水量에 큰 영향을 미친다고 할 수 있다. 대체적으로 給水量豫測에 있어서 回歸模型은 자료의 수집이 많을 때 훨씬 더 정확한 豫測을 할 수 있다. 給水量과 經年變化의 月別 相關係數는 0.8~0.9정도로 아주 높은 相關性을 보였다. 시계열 구간을 月別로 하였을 때의 給水量과 각 因子의 相關係數는 Table 2와 같다.

4.2 回歸模型式

多重回歸模型式의 主要因子에 대한 특성과 인자 상호간의 相關分析을 통하여 模型式을 제안하고 模型式을 이루는 因子와 模型式의 係數를 결정하였다. 模型式에 사용된 主要因子는 9개의 變數로서

給水量에 영향을 주지 않는 變數나, 說明變數 상호간의 相關度가 높은 變數를 제외하는 조합으로 4개의 回歸模型을 설정하고 이 지역에 가장 적합한 回歸模型을 도출하고자 하였다. 4개의 回歸模型式은 Table 3과 같이 구성하였다. Table 4는 시계열자료를 월별구간으로 구분한 모형 4에 대한 회귀계수이다.

4.3 多重回歸模型의 適合性 檢定

給水量 豫測을 위한 回歸模型式의 適合性 검토는 實測値와 豫測値를 비교하고 統計的으로 有意한지 여부를 檢定하기로 한다. 자료의 시계열구간은 月別, 季節別 및 年別期間 단위로 하였으므로 시계열기간 단위가 模型에 미치는 영향과 回歸模型式에 대한 適合性을 검토하여야 한다. 시계열구간에 따른 模型의 정도를 비교하기 위하여 修正決定係數 (adjusted coefficient of determination)와 模型別 誤差率을 살펴보았다.

시계열구간을 月別로 나누어 回歸模型을 구성할 때의 修正決定係數를 계산하여 정리하면 Table 5와 같다. 月別區間으로 回歸模型을 구성하였을 때, 修正決定係數의 최저값은 6월중에 88.9%로 나타나며 최고 94.8%로 7월 중에 나타나고 있다.

시계열구간을 季節別 및 年別로 나누어 回歸模型

을 구성할 때의 修正決定係數를 계산하면 Table 6에 나타난 바와 같다. 季節別로 回歸模型을 구성할 때에는 修正決定係數의 최저값은 83.3%이며 겨울철의 모형 4에서 나타난다. 또한 최대값은 92.8%로 가을철에 대하여 模型 2, 3, 4에서 나타나고 있다. 年別區間의 修正決定係數의 최대값 및 최소값은 각각 85.5%, 84.7%이다.

月別, 季節別, 年別에 따라 구성된 回歸模型에 대한 修正決定係數를 비교하면 回歸模型의 정확도는 月別, 季節別, 年別 단위로 구성된 模型式의 순서임을 알 수 있다. 따라서 給水量 豫測에 관한 回歸模型은 月別單位로 구성하는 것이 바람직하다고 판단된다.

回歸模型의 誤差率을 검토하기 위해 回歸模型을 月別, 季節別 및 年別單位로 구성하였을 때의 誤差率을 비교하였다. Table 7은 回歸模型을 月別單位로 구성할 때 발생하는 回歸模型의 誤差率이며 Table 8은 回歸模型을 季節別 및 年別單位로 구성할 때 발생하는 回歸模型의 誤差率이다. Table 7과 8에 나타난 誤差率을 비교하면 시계열구간을 月別單位로 구분하여 回歸模型을 구하는 것이 誤差率을 줄일 수 있다는 것을 확인할 수 있다. 따라서 修正決定係數에 이어서 誤差率의 비교에 있어서도 回歸模型은 月別單位로 구성하는 것이 정도가 좋다는 것을 확인 하였다.

Table 5. The Adjusted Coefficients of Determination of the Regression Model Obtained Using the Monthly Data for the Years of 1990 to 1994. (Units: %)

Category	Months	Model 1	Model 2	Model 3	Model 4
'90 - '94	Jan.	93.4	93.4	93.8	93.0
'90 - '94	Feb.	90.4	90.2	90.3	90.3
'90 - '94	March	91.3	91.3	90.9	90.8
'90 - '94	April	91.0	91.0	91.8	90.9
'90 - '94	May	92.2	92.7	92.6	92.6
'90 - '94	June	89.2	89.1	89.0	88.9
'90 - '94	July	94.8	94.5	94.2	94.0
'90 - '94	Aug.	91.4	91.6	91.3	91.3
'90 - '94	Sept.	94.4	94.4	94.3	94.3
'90 - '94	Oct.	93.8	93.7	93.8	93.8
'90 - '94	Nov.	91.4	91.4	91.5	91.4
'90 - '94	Dec.	91.4	91.4	91.5	91.4

Table 6. The Adjusted Coefficients of Determination of the Regression Model Obtained Using the Seasonal and Annual Data for the Year of 1990 to 1994. (Units : %)

Seasons and Years	Model No.1	Model No.2	Model No.3	Model No.4
Spring	91.0	90.6	90.3	90.4
Summer	87.1	85.8	85.8	85.4
Fall	92.8	92.4	92.4	92.6
Winter	84.7	84.3	84.0	83.3
Years	85.5	85.4	84.8	84.7

Table 7. Error Rate of the Regression Model Obtained Using the Monthly Data for the Years of 1990 to 1994. (Units : %)

Mon.	Model No. 1			Model No. 2			Model No. 3			Model No. 4		
	Max	Min	Ave	Max	Min	Ave	Max	Min	Ave	Max	Min	Ave
Jan	7.151	0.016	2.634	7.311	0.018	2.604	8.764	0.122	2.208	8.563	0.016	2.201
Feb	8.231	0.010	2.156	8.700	0.037	2.182	7.915	0.092	2.209	7.534	0.046	2.234
March	5.561	0.053	1.886	5.464	0.036	1.884	5.836	0.056	1.941	5.794	0.018	1.944
April	6.124	0.058	0.939	5.915	0.018	1.989	5.962	0.014	1.967	5.868	0.001	1.984
May	4.722	0.044	1.976	4.772	0.005	1.630	4.927	0.005	1.630	4.884	0.010	1.621
June	7.644	0.640	2.078	6.309	0.019	1.775	6.279	0.013	1.792	6.258	0.006	1.791
July	9.671	0.026	2.242	9.592	0.091	2.370	9.344	0.044	2.378	9.280	0.040	2.426
Aug.	6.230	0.021	1.896	6.256	0.064	1.896	6.026	0.009	1.924	5.892	0.047	1.912
Sept.	7.272	0.062	1.872	7.289	0.045	1.870	7.591	0.097	1.853	7.413	0.020	1.869
Oct.	7.632	0.004	1.825	7.872	0.010	1.891	8.383	0.014	1.991	8.142	0.015	1.867
Nov.	8.186	0.013	1.681	7.896	0.013	1.671	8.344	0.002	1.699	8.484	0.003	1.697
Dec.	8.749	0.011	1.835	8.415	0.012	1.819	8.741	0.019	1.847	8.603	0.010	1.839

$$\text{오차율} = \left| \frac{\text{실제급수량} - \text{예측급수량}}{\text{실제급수량}} \right| \times 100\%$$

Table 8. Error Rate of the Regression Model Obtained Using the Seasonal Data for the Years of 1990 to 1994. (Units: %)

Season and Years	Model No. 1			Model No. 2			Model No. 3			Model No. 4		
	Max	Min	Ave	Max	Min	Ave	Max	Min	Ave	Max	Min	Ave
Spring	8.002	0.005	2.318	7.902	0.005	2.039	7.663	0.025	2.074	7.763	0.006	2.061
Summer	13.929	0.019	2.895	12.545	0.029	3.081	16.223	0.026	3.854	12.673	0.036	3.134
Fall	13.646	0.061	2.075	14.925	0.069	2.083	15.185	0.027	2.113	14.690	0.050	2.053
Winter	14.523	0.030	2.606	15.384	0.036	2.651	15.678	0.001	2.679	10.416	0.048	2.681
Years	13.192	0.097	2.693	13.149	0.001	2.718	12.964	0.097	2.706	12.946	0.004	2.793

Table 7에 정리된 模型別 誤差率을 비교하여 보면 4개의 模型別 誤差率은 대동소이 한 것으로 판단되었다. 最大誤差率은 9.67%로 模型 1에서 발생했으며 시기적으로 7월에 해당한다. Table 7의

각 月別 모형에 따른 最大誤差率을 보면, 1월, 3월, 4월, 9월, 10월, 11월은 모형 1이 우수하지만, 2월, 4월, 6월, 8월, 12월은 모형 4가 모형 1보다 最大誤差率面에서 더 정확한 것을 알 수 있

Table 9. Analysis of Variance.

Model No.	Source	Degree of Freedom	Sum of Squares	Mean Squares	F	Critical Value(F)	Significance Level(α)
1	Regression	8	62234963968	7779370496	107.4	2.66	0.01
	Residual	115	8327048192	72409112			
	Total	123	70562013184				
2	Regression	6	62123278336	10353880064	143.55	2.96	0.01
	Residual	117	8438731264	72125919			
	Total	123	70562013184				
3	Regression	5	61318377472	12263675904	156.55	3.17	0.01
	Residual	118	9243634688	78335888			
	Total	123	70562013184				
4	Regression	5	61392445440	12278489088	158.01	3.17	0.01
	Residual	118	9169566720	77708192			
	Total	123	70562013184				

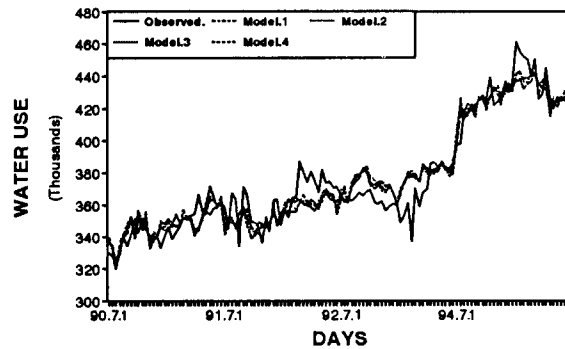


Fig. 2. The Comparison between the Measured and the Predicted Municipal Water Use for the Month of July from 1990 to 1994.

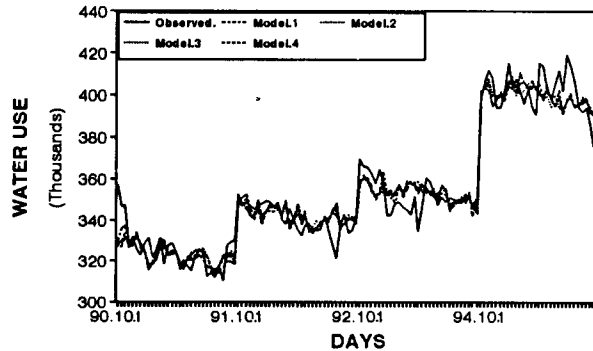


Fig. 3. The Comparison between the Measured and the Predicted Municipal Water Use for the Month of October from 1990 to 1994.

다. 平均誤差率을 보면, 模型 1 및 模型 4 모두 거의 비슷한 誤差率을 보이고 있다. 따라서 모든 變數를 고려한 模型 1보다, 給水人口, 最高氣溫, 降雨量, 日照時間, 經年變化만을 고려하여 모형을 구성하는 편이 더 實用的이라고 말할 수 있다.

統計學的인 檢定方式은 계산된 回歸模型이 統計的으로 有意한가(statistical significance)를 檢定하는 것이다. 適合性 檢定은 分散分析(analysis of variance)의 F 값을 계산한 후 계산된 F 값에 대한 분자와 분모의 自由度에 대응하는 臨界值가 有意水準에 적합한지를 판단하는 방식이다.

月別 模型式 1의 경우는 채택한 모든 變數를 포함하여 높은 相關性을 나타내고 있고, 多重共線性(multicollinearity)의 문제가 제기될 수 있으며, 說明變數가 많아 실제로 적용하기 어렵다. 따라서 모든 月別模型에 대해 分散分析을 수행하고 그 중에서 1월달에 해당되는 모형 1, 2, 3, 4에 대한 F 검정의 적용결과를 Table 9에 나타냈다.

실제로 F-檢定の 棄却值(critical value)는 $F_{(8, 115, 0.001)}$ 인데, $F=107.44$ 이므로 $F=107.44 > 2.66$ 로 도출한 回歸式은 給水量을 有意하게 설명하는 것으로 나타났다. 따라서 月別模型式 1, 2, 3, 4 모두 적합한 모형으로 판단되었다.

Fig. 2는 7월의 실측된 給水量과 모형에 의한 豫測值를 보여주고 있다. 모형간의 誤差率을 비교하면 最大誤差率은 모형 1에서 9.67%로 최대, 모형 4에서 9.27%로 최소, 平均誤差率에서는 모형 4보다 모형 1에서 최소를 보이고 있지만 그 차이는 미미하다. 따라서 모든 變數를 고려한 모형 1과 모형 4, 두 모형은 거의 비슷한 誤差率을 보이고 있기 때문에 모형 4가 더 실용적이라고 할 수 있다.

Fig. 3은 10월의 給水量과 豫測值를 보여주고 있다. 模型간의 誤差率은 비슷하나 最大誤差率은 模型 3에서 발생하였으며 最大誤差率은 8.3%로서 타 模型에 비하여 0.5% 정도 誤차가 큰 것으로 나타났다. 10월달 역시 平均誤差率은 2% 이내로 나타나고 있으며 F 檢定結果 回歸模型式이 有意水準 1% 범위 내에서 적합한 것으로 판정되었다.

앞에서 검토한 4, 10월을 제외한 나머지 1~3, 5~9, 11~12월에 대하여 조사한 결과, 나머지 달

에 대해서도 平均誤差率은 2% 이내이며 最大誤差率은 6~9% 이내에 들어가는 模型으로 판정이 되었다.

本 模型은 給水量이 急增하거나 急減할 때 誤差率이 커진다는 단점을 안고 있으나 誤差率이 최대 로 발생한 7월의 경우에도 模型의 平均誤差率이 2.5%내외, 最大誤差率이 10%내외이었다. 겨울철 模型은 季節別 模型 중에서 誤差가 가장 크게 발생하고 있는데 여름철 模型에 비하여 模型 1은 최대 0.594%, 模型 2는 2.83%, 模型 3과 模型 4에서는 각각 여름철의 誤差가 0.5%, 1.2% 정도 크며, 平均誤差率은 0.3~1.2% 정도 여름철의 誤差가 크다. 一日 給水量 豫測에 있어서 최대 10% 오차율은 2.4시간의 給水量에 해당하는 誤差로서 시설을 운영하는 측면에서 配水池容量에 건주어 비교하여 보면 수용할 만한 誤差率이라고 할 수 있다.

5. 結 論

本 研究에서는 光州廣域市를 대상으로 上水道 一日 給水量資料, 人口資料 및 氣候資料를 이용하여 回歸分析을 실시하였다. 또한 日別로 變化되는 給水量을 豫測하기 위한 回歸模型式을 제안하였으며, 다음과 같은 結論을 얻었다.

1. 上水道 一日 給水量을 豫測하는 主要因子로 人口數와 經年變化 및 氣候因子를 이용하여 線形 多重回歸模型을 구성한 결과, 模型의 平均誤差率은 평균 3% 이내가 되며 最大誤差率은 10% 이내로서 비교적 좋은 정확도를 갖는 模型으로 판명되었다.
2. 上水道 一日 給水量을 豫測하기 위한 回歸模型의 시계열구간을 年別, 月別, 季節別區間으로 분류하여 예측한 결과 年別, 季節別區間 보다 月別區間の 정도가 높아 一日 給水量을 月別區間으로 豫測하는 것이 적절하다고 판단 되었다.
3. 上水道 一日 給水量을 豫測하는 回歸模型의 氣候條件을 나타내는 7개의 氣候因子인 一日 平均기온, 일일 최고기온, 상대습도, 강우량, 운량, 일조시간, 일사량중에서 일일 최고기온, 일조시간, 강우량은 중요한 說明變數로 파악 되었으며, 이들 因子와 給水人口, 經年變化를 고려

한 모델은 비교적 정도가 높은 모델으로 판단되었다.

4. 본 연구에서 제시한 모델은 上水道施設の 운영 시 1日 計劃給水量을 추정하는 데 유용하게 사용할 수 있으나 生産量과 급·배수지역의 需要量을 실시간으로 비교하여 單位 時間別로 요구되는 生産量을 추정하는 데는 한계가 있으므로 이에 대한 추가적인 연구가 필요하다고 사료된다.

參 考 文 獻

- 강변서 (1993). 行列과 SPSS/PC+ 이용 多變量 統計 分析. 학현사.
- 光州 統計年報. (1982-1994). 광주광역시.
- 氣象年報. (1982-1994). 기상청.
- 氣象月報. (1982-1994). 기상청.
- 김수원 (1985). “都市用水 計劃을 위한 물 需要 豫測에 관한 研究,” 박사학위 논문, 영남대학교.
- 안상형, 이명호 (1993). 現代統計學. 학현사.
- 이경훈, 박성천, 문병석 (1994a). “上水道의 給水量 및 配水池 容量變化에 관한 研究.” 대한상하수도학회지, 제9권, 제2호, pp. 118-126.
- 이경훈, 이삼노, 문병석 (1994b). “上水道 1日 給水量의 時間的 變化의 特性에 관한 研究.” 한국수문학회지, 제27권, 제2호, pp. 135-143.
- 이경훈, 이삼노, 문병석 (1994c). “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.
- Peak 용수수요 분산을 고려한 광역상수도 용수요율 체계의 개선. (1992). 한국수자원공사.
- 현인환, 목동우 (1995). “上水 給水量의 短期 豫測.” 대한상하수도학회지, 제9권, 제1호, pp. 96-106.
- Frankeland, R.J., and Shouvanavirakul, P. (1973). “Water consumption in small communities of northeast Thailand.” *Water Resources Research*, Vol. 9., No. 5, pp. 1196-1207.
- Howe C.W., and Linaweaver, F.P. (1967). “The impact of price on residential water demand and it's relation to system design and price structure.” *Water Resources Research*, Vol. 14, No. 4, pp. 13-32.
- 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: 1995년 7월 10일)