

세탁기 및 화장실 용수 수요량에 대한 예측모형 연구

명성민*

A Study on Prediction Model for Laundry and Toilet Water-use demand

Sung-Min Myoung*

요약 본 연구는 5년간의 세탁기 및 화장실 용수의 실측 자료를 기반으로 표본가구의 가구 및 주택, 월 특성들을 나타내는 변수들을 조사하여 위의 두 가지 용수에 대한 수요예측모형을 개발하는 것이다. 그러나 반응변수인 세탁기 및 화장실 용수의 분포를 확인한 결과 양의 왜도, 즉, 왼쪽으로 치우친 형태로 정규분포를 따르지 않기 때문에 다중회귀모형 적용 시 추정치가 편되는 문제가 있다. 따라서 이에 대한 대안으로 세탁기 및 화장실 용수 수요 예측모형을 와이불 및 대수정규회귀모형으로 가정하고, 3가지 모형을 적용하여 최적모형을 제시하고 이에 대한 해석을 제시하였다. 그 결과 와이불 분포를 가정한 회귀모형이 가장 적합하다고 나타났으며, 이에 대한 각 용수별 예측모형을 제시하고 해석하였다. 그 결과 두 용수 모두 공통적으로 실거주인원이 용수사용에 가장 큰 영향을 주는 인자로 나타났다. 분석결과를 토대로 세탁기 및 화장실 용수의 수요를 예측 시 실거주인원에 대한 고려가 필요하다는 점을 시사하며, 예측모형을 통한 관리부서에서 장기적으로 물 수요관리에 대한 정책 수립, 수도 관련 시설 규격 및 기자재 결정 등의 기초자료로 활용 할 수 있을 것이라 판단된다. 추후 연구에서는 1인 가구 등과 같은 실태를 반영한 독립변수들을 고려한 실측연구 등이 필요할 것이다.

Abstract This study develops a prediction model for toilet and laundry water end-uses based on surveyed data which measured housing and household characteristics of 140 households over 5 years in Korea. Classical regression model assuming a normal distribution was not appropriate and estimated parameters were biased, because the distribution of measured water-uses was left-skewed. As an alternative to this problem, we considered the distribution of weibull and lognormal for each water-uses, and three regression models were compared using log-likelihood and scale parameter. As a result, weibull regression were chosen to be appropriate for both water-uses and also presented the factors that affect each water-use. This results expect that an insight is provided on water resources utilization and theoretical support role for effective water resource management.

Key Words : Water Conservation, Prediction model, Weibull distribution, Lognormal distribution, Regression model

1. 서론

2018년 4월에 환경부에서 발표된 상수도 통계에서 2016년 업종별 물 사용량은 가정용이 63.6% (3,413백만^m)으로 가장 많고, 그 다음으로 영업용 28.3% (1,519백만^m), 공업용 2.8% (150백만^m), 욕탕용 1.7% (90백만^m), 기타 0.4% (19백만^m) 순으

로 나타났다. 이는 2015년 통계에서 발표된 가정용 물 사용량 64.0% (3,444백만^m)보다는 약간 감소하였으나, 가정용수가 가장 많은 비중을 차지한다는 점에서는 이견이 없다 [1]. 이러한 상황에서 가정용수에 대한 영향인자 분석 및 특성파악을 위한 많은 연구들이 진행되어져 왔는데, 일반적으로 가정용수에 영향을 미치는 요인으로써 날씨, 온도 등의 외부적

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A1B03029018)

*Corresponding Author : Department of Health Administration, Jungwon University (smmyoung@jwu.ac.kr)

Received June 26, 2019

Revised August 03, 2019

Accepted August 03, 2019

요인과 주택구조와 주거형태와 같은 내부적 요인으로 구분된다 [2].

이러한 영향인자 분석을 통하여 용수의 수요량 예측 모형을 제시하는 최초의 방법으로 미국 Oklahoma City와 Oklahoma주 Tulsa에서 1961년부터 20년간 실측한 1인당 물 소비량을 조사하여 물의 가격, 1인당 소득, 강수량, 기온, 인구1000명당 가구 수 등을 영향 변수로 고려하여 다중회귀분석을 수행하였다 [2]. 이와 같이 가정용수에 대한 사용량의 실측을 통하여, 정책제시, 물 손실관리/수요예측 및 도시 수도 시스템의 최적 운영 등과 같은 추후 방향설정 등으로 활용하고 있다. 이에 대한 국내의 연구경향을 보면 다음과 같다.

기존 해외 연구에서는 미국 수도협회(American Water Works Association)에서 1999~2000년까지 1년 동안 수도요금 청구서를 조사하고 12개 지역별 100여개의 표본가구를 대상으로 물 소비패턴을 실측조사하였다. 또한 미국의 덴버 수도국에서는 1995년~2005년간 단독주택을 대상으로 용수 실측조사를 수행하여 경향성을 파악하였다 [3-4]. 또한 호주 Melbourne지역의 수요량 예측을 위하여 추세, 계절성, 기후상관 및 자기상관의 4개의 요인을 고려한 시계열 모형을 제시하였고, 1년간의 자료를 가지고 개발된 예측모형을 검증한 연구도 있으며, 다른 연구자는 물 수요에 대한 예측을 3단계(rising, oscillating, falling)로 식별하기 위한 패턴인식 접근방법을 제시하였다 [5-6]. Adamowski 등(2010)은 주간 물 수요예측모형으로 인공신경망(ANN: Artificial Neural Network) 모형과 다중선형회귀분석과 비교하여, 인공신경망 모형이 주간 물 수요에 대한 예측의 우수성을 확인하였다. 또한 Nasserri 등(2011)은 1992~2002년까지 10년간이란 Tehran의 월간 물 사용량을 기반으로 확장된 칼만 필터(extended Kalman filter)와 유전자 프로그래밍(genetic programming)을 결합한 하이브리드 모형을 개발하고, 이에 대한 예측모형의 정밀성을 입증하였으며, Brentan 등(2017)은 SVM(support vector machine)과 푸리에급수(Fourier series)를 고려한 결합방법을 이용한 결합모형을 제안하였다 [7-9]. 최근에 Duerr 등(2018)은 미국 플로리다 남서부지역의 Tempa Bay Water에 의해 수집된 빅데이터를 이용하

여 공간통계모형(spatiotemporal statistical model)과 기계학습 알고리즘을 이용하여 월간 물 사용에 대한 단기 예측모형을 구축한 연구도 있다 [10].

우리나라에서는 한국수자원공사에서 연구한 지역별 140여개 표본조사가구를 대상으로 실측한 용도별 가정용수를 대상으로 집단별 평균검정, 상관관계를 수행하였고, 시계열 분석을 통하여 계절요인의 존재성을 확인하였다. 또한 비가정용수에 대하여 업종을 세분화하여 이에 대한 사용량 원단위를 도출하고 이에 대한 관리방안을 제안한 연구가 있다. 또한 동태적 다지역 수자원(recursive multi-region water) CGE 모형을 개발하여 지역별 용수수요예측을 나타낸 연구도 있으며, 최근의 연구에서는 강수량, 기온, 풍속과 같은 기상 자료를 기반으로 회귀모형을 제시하여 농업용저수지 저수량 예측모형을 제시하고, ROC곡선 등을 통해 모형의 적합도를 확인한 것도 있다 [11-14].

본 연구는 가정용수 중 화장실과 세탁용수에 초점을 두어 실측한 자료를 대상으로 내부적, 외부적 요인을 설명변수로 설정하고, 1인당 물사용량을 반응변수로 하여 수요예측모형을 제시하고자 한다. 화장실 용수 및 세탁기 용수에 초점을 둔 이유는 기존 연구에서 가정용수의 용도별 용수 사용량에서 화장실과 세탁기 용수가 가장 높았기 때문이다 [11].

측정된 용수자료의 분포를 확인한 결과, 왼쪽으로 치우치며 우측 긴 꼬리를 가지는 형태로 나타나, 좌우 대칭을 가지는 정규분포 가정을 만족하지 않았다. 그러므로 기존의 회귀분석 적용 시 설명력이 낮게 나타나며, 추정치가 편의된(biased) 값으로 나타나게 된다.

따라서 본 연구에서는 화장실 용수 및 세탁용수의 1인당 물 사용량에 대한 회귀분석 적용시, 보편적으로 사용되어온 정규분포를 적용하는 것이 아닌, 와이블 분포 또는 대수정규분포를 사용한 회귀모형을 적용하고, 이를 비교하여 가장 적절한 분포모형을 결정하고, 결정된 모형을 가지고 예측모형을 제시하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 화장실 용수 및 세탁용수 자료의 취득방법 및 전처리(pre-processing)를 설명하고, 대수정규회귀 및 와이블 회귀모형에 대한 이론적 고찰을 제시한다. 3 장에서는 전처리된 자료를 기반으로 하여 가능한 모형에 대한 적합도를 제시하고, 이를

비교하여 최적 모형을 결정하여 데이터에 적용한 결과를 제시하고 그 의미를 설명한다. 마지막 4장에서는 연구결과를 요약하고 향후 연구과제에 대하여 논의한다.

2. 분석 자료 및 방법

2.1 자료

본 연구에서 사용된 자료는 55개 시·군을 대상으로 하여 140개 가구를 표본으로 선정하여 한국수자원공사에서 5년간 용도별 유량을 실측한 자료이다. 본 연구에서는 가정용수의 용도 중 화장실 용수와 세탁기 용수에 대하여 고려한다. 유량실측 방법은 그림 1과 같다 [15]. 먼저 각 가정 내 수도꼭지에 전자식 유량계를 설치하고, 자료저장 및 취득한 자료를 무선송신이 가능하도록 유량계에 로거를 부착한 모니터링 시스템을 제작하여 설치하였으며, CDMA 방식의 무선신호를 가정 내 각각의 전자식 유량계에서 전송하고 중계기를 거쳐 유량자료를 수집하여 해당 서버에 저장하는 방식을 이용하였다 [13,14].

2.2 자료의 전처리

위에서 제시한 바와 같이 수집된 용도별 유량자료를 그대로 이용할 수 없기 때문에 자료에 대한 전처리를 실시하였다. 주요 전처리 방법으로 이상치(outlier)를 제거하였다. 이상치 제거는 탐색적 자료분석 (exploratory data analysis)에서 주로 이용하는 개념으로 외부상한(upper outer fence)을 이용하였다.

세부적으로는 전체 가구를 대상으로 용도별 유량측정자료는 화장실 용수 122,730개, 세탁기 용수 120,422개였으며, 이를 대상으로 전처리를 실시하였다. 첫 번째 단계로, 측정된 용수값이 음수로 나타난 경우로서, 이는 기계의 오류로 인한 경우라 판단하여 제거하였다.

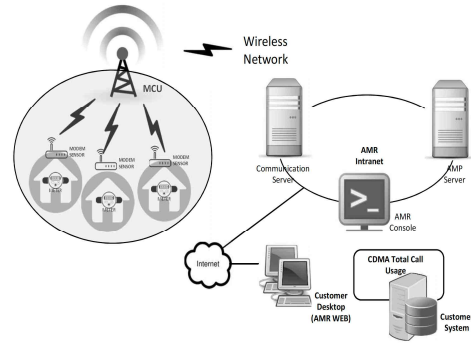


그림 1. 가정용수 취득과정
Fig. 1. Data collection procedure of end-use water in household

두 번째 단계로 upper outer fence를 초과하는 용수값이 나타난 경우를 제외하였다. 자료의 특성상 낮게 나타나는 자료는 존재하지 않았기 때문에, 상한값만을 고려하였으며, 이를 초과하는 자료의 경우는 사용이 어려울 정도로 과다하게 사용되었다고 판단하였기 때문이다. 마지막 단계로, 필터링된 자료를 대상으로 가구별 1인당 물 사용량 (pcd: liter per capita day)를 계산하였으며, 이를 종속변수로 설정하였다. 그러나, 종속변수인 일별 1인당 물 사용량이 개인별로 차이가 심하여 오차분산이 높게 나타나서, 이를 보완하기 위하여 월 평균 1인당 물 사용량으로 변환하였다.

확정된 화장실 및 세탁기 용수의 월 1인당 물 사용량에 대한 시계열 도표는 그림 2와 같다. 대체적으로 화장실 용수가 세탁용수보다 물 사용량이 더 높으며, 월별요인이 존재함을 확인할 수 있었다.

2.3 용수사용량에 대한 와이블 및 대수정규회귀 모형

본 절에서는 대수정규회귀 및 와이블 회귀모형에 대하여 기술한다. 각 용도별 용수사용량을 t 로 정의했을 때, t 는 0이 아닌 양의 값을 항상 가지기 때문에, 일반적인 선형회귀모형을 이용하여 t 를 모형화 하는 것은 적절하지 않다. 이를 해결하기 위한 방안으로 가장 명확한 방법으로는 $y = \ln t$ 로 변환하는 로그변환을 수행하는 것이다.

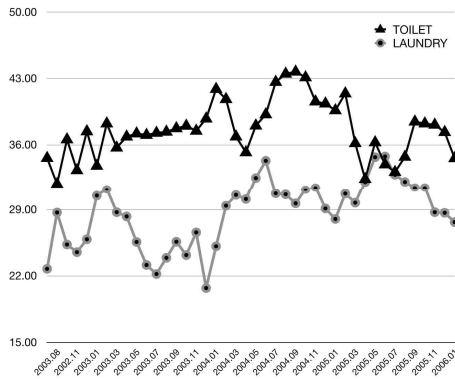


그림 2. 세탁기 및 화장실 용수자료에 대한 시계열도표
 Fig. 2. Time series graph for monthly laundry and toilet water-uses dataset

이렇게 변환한 경우 일반적으로 와이블 분포 및 대수정규분포를 따른다고 알려져 있다[15]. 와이블 회귀 모형 및 대수정규회귀모형에 대한 세부 내용은 다음과 같다.

첫 번째, 대수정규 밀도함수(lognormal density function)는 다음과 같다.

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} t^{-1} \exp\left(-\frac{(\ln t - \mu)^2}{2\sigma^2}\right), t > 0 \dots\dots\dots\text{식(1)}$$

두 번째로, 와이블 밀도함수는 다음과 같다. 여기서 λ 는 척도모수 (scale parameter)이며, φ 는 형상모수 (shape parameter)이다.

$$f(t) = \frac{\varphi}{\lambda} \left(\frac{t}{\lambda}\right)^{\varphi-1} \exp\left[-\left(\frac{t}{\lambda}\right)^\varphi\right], t > 0 \dots\dots\text{식(2)}$$

용수사용량 t 가 대수정규분포를 따르는 경우 $y = \ln t$ 는 정규분포 $N(\mu, \sigma^2)$ 를 따르며, 와이블 분포를 따르는 경우 다음의 극단값 분포(extreme distribution)을 따르는 밀도함수를 갖는다. 여기서 $\mu = -\ln \lambda$ 는 위치모수이며, $\sigma = \varphi^{-1}$ 는 척도모수이다.

$$f(y) = \frac{1}{\sigma} \exp\left[\frac{y - \mu}{\sigma} - \exp\left(\frac{y - \mu}{\sigma}\right)\right] \dots\dots\text{식(3)}$$

$i = 1, 2, \dots, n$ 인 경우 로그변환된 y 를 적용한 선형 회귀모형은 다음과 같다.

$$y_i = \ln(t_i) = x_i^T \beta + \sigma \epsilon_i \dots\dots\dots\text{식(4)}$$

$$= \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \sigma \epsilon_i$$

여기서 $x_i^T = (1, x_{i1}, \dots, x_{ip})$ 는 공변량 벡터들이며, $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$ 는 회귀계수, σ 는 척도모수이다. ϵ_i 는 오차항이며 와이블분포를 따르는 경우 극단값분포를 따르며, 대수정규분포인 경우 표준정규분포 $N(0,1)$ 를 따른다.

와이블 회귀모형 및 대수정규회귀모형의 모수추정 방법은 최대우도추정법(maximum likelihood estimator: MLE)을 사용한다. 대수정규회귀모형에서 모수들에 대한 우도함수는 다음과 같다.

$$L(\beta, \sigma^2) = \prod (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{[(y_i - x_i^T \beta)/\sigma]^2}{2}\right) \dots\dots\dots\text{식(5)}$$

또한 와이블 회귀모형의 경우 우도함수는 아래와 같다.

$$L(\beta, \sigma^2) = \prod \left(\frac{1}{\sigma}\right) \exp\left\{\left[\frac{(y_i - x_i^T \beta)}{\sigma}\right] - \exp\left[\frac{(y_i - x_i^T \beta)}{\sigma}\right]\right\} \dots\dots\text{식(6)}$$

우도함수를 얻게 되면 모수들에 대하여 우도함수를 최대화하여 추정치를 얻게 된다. 추정치의 검정을 위하여 표준오차를 계산해야 하는데, 이는 피셔의 정보행렬 I 또는 관찰된 정보행렬 I_0 를 다음과 같이 정의하여 얻을 수 있다. 여기서 $\hat{\theta}$ 는 최대우도추정치이다. 또한 i 번째 모수의 유의성을 검정하기 위해 우도비검정(likelihood ratio test)를 적용할 수 있다.

$$I(\hat{\theta}) = E\left(\frac{-\partial^2 L(\hat{\theta})}{\partial \theta_i \partial \theta_j}\right) \Bigg|_{\theta = \hat{\theta}}$$

$$I_0(\hat{\theta}) = E \left(\frac{-\partial^2 L(\hat{\theta})}{\partial \theta_i \partial \theta_j} \right) \Bigg|_{\theta = \hat{\theta}} \dots\dots\dots \text{식(7)}$$

귀무가설 $H_0 : \theta_i = 0$ 를 검정하기 위한 검정통계량 Γ 은 다음과 같으며, 자유도 1을 가지는 카이제곱분포를 따른다.

$$\Gamma = -2 \ln \frac{L(\hat{\theta})_{\{-i\}}}{L(\hat{\theta})} \dots\dots\dots \text{식(8)}$$

위 식과 같이 최대우도추정치값의 존재여부에 대한 정확한 조건은 Silvapulle and Burrige(1986)에 의하여 증명 되었으며, 추정가능도의 문제는 자료의 개수와 모수의 수가 거의 같은 경우 발생할 수 있다고 주장 하였다. 그러나, 비록 최대우도추정치가 무한하게 나타난다고 하더라도, 우도함수 값이 잘 정의되어있기 때문에 우도비 검정을 사용할 수 있다고 Clarkson and Jennrich (1991)가 입증하였다 [15].

2.4 통계분석 방법

본 연구에서 주 분석변수인 화장실 용수 및 세탁기 용수의 월별 사용량에 대한 분포는 그림 3과 같다. 분포의 모양이 명확하게 왼쪽으로 치우치며 우측 긴 꼬리를 가지는 형태이므로 정규분포 가정을 만족한다고 주장하는 것은 어려울 것이라 판단하였다. 두 용수의 월별 사용량에 대한 수요예측모형을 설정하기 위한 종속변수는 두 용수별 월별 사용량으로 설정하였으며, 독립변수는 기존 Cochran and Cotton이 제안했던 사회적, 환경적 요인을 고려하였다. 세부적으로 본 연구에서는 기존의 Kim 등(2007)의 연구에서 제안했던 변수를 독립변수로 고려하였는데, 사회적 요인으로는 주택유형 (housing type), 실거주인원 (number of residents), 방 개수 (number of rooms) 및 건평 (floor space)을, 환경적 요인으로는 월별 효과 (monthly effect)를 설정하였다 [11].

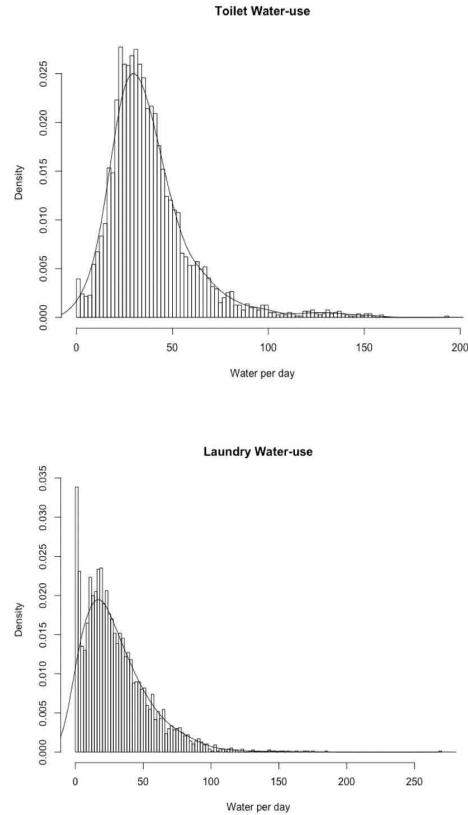


그림 3. 세탁기 및 화장실 용수자료의 분포
Fig. 3. Distributions of toilet and laundry water-uses

통계 분석 프로그램은 R-package 3.5.3 및 Rstudio 1.14을 이용하였다. 통계검정시 유의수준은 5%로 설정하였으며, p-value가 0.05보다 작은 경우 유의하다고 판정하였다.

3. 분석결과

선정된 독립변수들에 대한 화장실 및 세탁기 1인당 용수 사용량에 대한 기술통계 결과는 표 1과 같다. 화장실 용수 및 욕조 용수의 1인당 사용량은 실거주인원 및 방의 개수와 건평 모두 양의 상관관계가 존재하였으며, 통계적으로 유의하게 나타났다. 주택유형의 경우에서도 분산분석 (analysis of variance)결과 통계적으로 유의하게 나타났으며, Scheffe 방법을 이용한 사후검정(post-doc test)결과 아파트 거주가 다른 거주

보다 용수사용량이 평균적으로 유의한 차이로 나타났다. 즉, APT 거주자가 단독주택 또는 맨션보다는 화장실 및 세탁기 용수를 더욱 많이 사용한다는 것이다.

표 1. 세탁기 및 화장실 용수자료의 기술통계량
Table 1. Descriptive statistics for laundry and toilet water-uses

Variable		Laundry	Toilet
number of residents [†]		0.3304*	0.1748*
number of room [‡]		0.0785*	0.1180*
floor space [†]		0.0342*	0.1242*
Housing Type ^{†,**,††}	Standalone	35.71±19.91	29.52±22.02
	APT	40.98±22.42	31.20±26.20
	APT home	35.99±23.02	29.72±20.75
Monthly Effect ^{†,**,††}	JAN	38.45±23.39	26.94±20.85
	FEB	39.81±22.84	28.17±22.59
	MAR	36.01±18.28	30.85±23.23
	APR	36.33±19.36	30.85±27.23
	MAY	37.36±20.32	31.64±25.61
	JUN	37.69±22.35	33.38±26.29
	JUL	39.58±23.01	34.39±26.78
	AUG	39.77±23.55	31.51±23.45
	SEP	39.30±23.00	31.02±24.41
	OCT	39.80±22.15	30.17±23.08
	NOV	37.45±21.52	29.62±23.20
	DEC	38.94±22.00	29.24±23.08

† calculated by correlation coefficient
‡ calculated by mean±sd
* p-value(0.05 by correlation analysis
** p-value(0.05 by analysis of variance

월별효과를 확인하기 위하여 분산분석 수행결과, 통계적으로 유의하게 나타났다. 선정된 독립변수와 종속변수에 대한 가능한 회귀모형은 정규분포, 와이블, 대수정규를 가정한 3개의 회귀모형으로 고려하고, 가정한 위 3개 회귀모형 중 가장 적합한 모형의 회귀모형을 선택하기 위하여 적합도 판정 지수를 이용하였다. 각 모형별 로그우도함수값 및 척도모수는 표 2와 같다.

모형의 적합성에 대한 판정하기 위한 기준으로 로그우도함수가 클수록, 척도모수는 작은 분포가 적합하다고 알려져 있다.

화장실 용수에 대한 1인당 사용량의 경우 가장 큰 로그우도함수는 대수정규분포를 가정한 경우이며, 척도모수의 경우 와이블 분포가 가장 작게 나타났다. 그러나, 화장실 용수의 로그우도함수값의 크기가 대수정규분포로 가정한 경우 (-3046.75)와 와이블 분포로 가정한

경우 (-3061.71)와 그리 많은 값의 차이가 존재하지 않았다. 따라서, 척도모수(0.48)가 명확하게 작게 나타난 것을 우선하여 와이블 분포로 모형을 선정하였다.

표 2. 용수자료의 적합도 검정
Table 2. Goodness of fit in water-uses

Distribution	Type	Laundry	Toilet
Normal	log-likelihood	-17555.85	-17402.75
	Scale Parameter	23.86	20.53
Log-normal	log-likelihood	-6032.11	-3046.75
	Scale Parameter	1.22	0.53
Weibull	log-likelihood	-5150.14	-3061.71
	Scale Parameter	0.79	0.48

세탁기 용수의 경우 와이블 분포가 가장 적합 (로그우도함수값 -5150.14, 척도모수 0.79)하다고 판정되었다. 선정된 독립변수와 용수별 1인당 물 사용량을 종속변수로 한 와이블 회귀모형 결과는 각각 표 3(화장실용수) 과 표 4(세탁기 용수)와 같다.

화장실 용수의 경우 실거주인원, 건평, 방 개수, 공동주택 대비 단독주택인 경우, 12월에 비하여 3~4월의 월효과가 통계적으로 유의하게 나타났다.

세부적인 해석은 다음과 같다. 먼저 실거주인원의 경우 1인당 화장실 용수 사용량의 로그값이 0.2101로 증가하는 형태로 나타났으며, 방의 개수인 경우에도 0.1866으로 증가하는 형태로 나타났다. 건평의 경우는 -0.0069로서 비록 통계적으로 유의하기는 하지만 거의 0에 가까우므로 미세한 감소추세를 가진다고 할 수 있다. 주택유형의 경우 가변수를 이용하여 공동주택(APT house)에 비해 단독주택은 1인당 화장실 용수 사용량의 로그값이 -0.0985 작게 나타나며, 반대로 공동주택에 비해 아파트는 0.0146 높게 나타난다는 것이다. 즉, 아파트, 공동주택, 단독주택 순으로 화장실의 1인당 용수 사용량이 나타난다는 것이다. 가변수로 설정한 월별 효과의 경우 12월에 비해 3월~4월에 사용량이 감소하며, 나머지 월효과는 미비함을 확인하였다. 설정된 와이블 회귀모형에 대한 형상모수는 2.0809, 척도모수는 0.4806으로 나타났다. 표준화 회귀계수를 통하여 확인한 결과 화장실 용수의 1인당 물 사용량에 가장 영향을 미치는 요인은 거주자 수(B=26.59)와 방

의 개수 (B=12.52)로 나타났다. 또한 계절 요인도 비표준화계수와 마찬가지로 3~4월이 가장 큰 연관성을 보였다. 각 독립변수들 간의 상관성은 존재하지 않았다.

표 3. 화장실 용수자료의 와이블 회귀 결과
Table 4. Weibull regression result in toilet water-use

Variable	Beta [†]	Std. Error	B [‡]	Chi-sq	
intercept	4.2600	0.0488	.	7618.7	
number of residents	0.2101	0.0079	26.59	699.99	
number of room	-0.0069	0.0007	-9.86	105.47	
floor space	0.1866	0.0149	12.52	156.78	
Housing Type	Standalone	-0.0985	0.0229	-4.30	18.44
	APT	0.0146	0.0213	0.69	0.47
	APT home	0	.	.	.
Monthly Effect	JAB	-0.0080	0.0344	-0.23	0.05
	FEB	0.0124	0.0346	0.36	0.13
	MAR	-0.1057	0.0374	-2.83	8.01
	APR	-0.0839	0.0374	-2.24	5.03
	MAY	-0.0566	0.0374	-1.51	2.29
	JUN	-0.0361	0.0376	-0.96	0.92
	JUL	0.0179	0.0376	0.48	0.23
	AUG	0.0237	0.0376	0.63	0.4
	SEP	0.0054	0.0369	0.15	0.02
	OCT	0.0116	0.0351	0.33	0.11
	NOV	-0.0403	0.0344	-1.17	1.37
	DEC	0	.	.	.

Scale parameter=0.4806, Shape parameter=2.0809
[†] Unstandardized regression estimate
[‡] Standardized regression estimate
 * statistically significant p-value<0.05

세탁기 용수에 대한 1인당 용수 사용량의 경우 실거주인원, 건평, 12월에 비하여 6~7월의 월효과가 통계적으로 유의하게 나타났다 (p-value<0.05).

화장실 용수의 결과와 비교해 볼 때, 방의개수와 주택유형이 세탁용수에 대한 1인당 용수 사용량의 경우에는 통계적 유의성이 존재하지 않았다. 실거주인원과 건평의 경우는 화장실 용수의 경우와 동일한 추세가 존재하였다.

그러나, 비록 통계적 유의성은 존재하지 않았어도 방의 개수의 경우에는 화장실 용수의 1인당 사용량과는 다르게 방의개수가 음의 관계를 가진다는 점이였다.

표 4. 세탁기 용수자료의 와이블 회귀 결과
Table 4. Weibull regression result in laundry water-use

Variable	Beta [†]	Std. Error	B [‡]	Chi-sq	
intercept	4.1122	0.081	.	2578.96	
number of residents	0.1291	0.0147	8.78	77.19	
number of room	0.0041	0.0012	3.42	11.9	
floor space	-0.033	0.022	-1.50	2.25	
Housing Type	Standalone	0.045	0.0403	1.12	1.24
	APT	0.0781	0.0364	2.15	4.61
	APT home	0	.	.	.
Monthly Effect	JAB	-0.0917	0.0582	-1.58	1.48
	FEB	-0.0352	0.0585	-0.60	0.36
	MAR	0.0532	0.0633	0.84	0.71
	APR	0.065	0.0633	1.03	1.06
	MAY	0.0819	0.0631	1.30	1.69
	JUN	0.1288	0.0632	2.04	4.15
	JUL	0.1506	0.0636	2.37	5.6
	AUG	0.0587	0.0633	0.93	0.86
	SEP	0.0518	0.0624	0.83	0.69
	OCT	0.0242	0.0594	0.41	0.17
	NOV	0.0086	0.0583	0.15	0.02
	DEC	0	.	.	.

Scale parameter=0.7943, Shape parameter=1.2589

[†] Unstandardized regression estimate
[‡] Standardized regression estimate
 * statistically significant p-value<0.05

주택유형의 경우에도 통계적 유의성은 존재하지 않았지만 공동주택에 비해 아파트, 단독주택이 더 크게 나타났다는 점이 화장실 용수의 1인당 사용량과는 다른 추세로 나타난다고 할 수 있다. 실거주인원의 경우 1인당 세탁기 용수 사용량의 로그값이 0.1291, 방의 개수는 0.0041, 건평의 경우 -0.033이었으며, 주택유형은 공동주택에 비해 단독주택 0.045, 아파트 0.0781로 나타났다.

월별효과와 계절 효과의 경우 12월에 비해 상대적으로 온도가 올라가는 6~7월에 유의성이 존재하였다. 비록 유의성은 존재하지 않았지만 이후 11월까지의 상대적으로 12월보다 용수사용량이 많았으며, 1월과 2월의 경우는 사용량이 작게 추정되어, 계절적 요인이 존재한다고 판단하였다. 설정된 와이블 회귀모형에 대한 형상모수는 1.2589, 척도모수는 0.7943으로 나타났다.

표준화계수를 통해 확인한 결과 세탁기 용수의 1인당 사용량과 가장 큰 연관정도를 보이는 요인은 거주자 수(B=8.78)와 건평(B=3.42)로 나타났다. 계절요인

의 경우에는 6~7월이 상대적으로 높은 표준화계수로 나타났다. 화장실 용수와 마찬가지로 독립변수들 간의 상관성은 존재하지 않았다.

4. 결론 및 논의

위에서 기술하였듯이 가정용수는 사회적, 환경적 요인 등에 의하여 영향을 받는다. 또한 Kim 등 [11]의 연구에서 가정용수의 용도별 물 사용비용을 조사한 결과 화장실, 세탁기, 싱크대, 욕조용수 순으로 나타나, 절수계획 수립 시 이를 통한 절수효과가 가장 높을 것이라 주장하였다.

본 연구는 실시간 용수 측정을 위하여 무선네트워크 통신 시설을 이용한 데이터 송/수신기법들을 활용하여 가정용수 중 화장실 용수와 세탁기 용수의 1인당 사용량에 대한 예측모형을 제시하고자 하였다. 화장실 및 세탁기 용수에 대한 분포의 특성이 항상 양수이고, 분포의 모양이 왼쪽으로 치우친 형태로 나타났기 때문에 고전적인 정규분포를 가정한 예측모형 제시가 어려울 것이라 판단하였다. 이에 대한 대안으로 예측모형에 대한 분포가정을 대수정규분포와 와이블분포로 설정하고, 이에 대한 최적의 분포를 설정하기 위해 위의 세 가지 분포를 가정한 예측모형을 로그우도함수와 척도모수를 이용하여 비교하였다. 그 결과 화장실 및 세탁기 용수 모두 와이블 분포를 가정한 회귀모형이 가장 적합한 모형으로 설정하였다. 각 용수별 예측모형을 제시하였으며, 화장실 용수에 대한 1인당 사용량의 경우 실거주인원, 건평, 방의 개수, 주택유형, 12월에 비하여 3~4월의 월효과가 영향을 미치는 주요한 인자로 해석되었으며, 세탁기 용수에 대한 1인당 사용량 경우 실거주인원, 건평, 12월에 비하여 6~7월의 월효과가 주요인자로 나타났다. 두 용수 모두 공통적으로 실거주인원이 용수사용에 가장 큰 영향을 주는 인자라고 발견되었다. 이는 향후 용수사용량 및 수요측면에서 실거주인수에 대한 고려가 필요하다는 것을 시사한다.

본 연구를 통하여 용수 수요관리 측면에서 이러한 화장실 및 세탁기 용수의 1인당 사용량과 패턴에 대한 예측모형의 제시는 수도 관련 정책 및 생활패턴의 변화를 이해하고 예측하는 자료로서의 가치가 있을 것이

라 기대된다. 또한 예측모형을 통한 관리부서에서 수요 예측, 수도관련 시설의 규격 및 기자재 결정 등에 활용할 수 있을 것이라 판단된다.

본 연구에서는 예측모형으로 회귀모형만을 이용하였으나, 추후 인공지능망 및 하이브리드 시계열 모형 등을 적용하여 이들 모형들을 비교하여 더욱 정밀한 예측모형을 개발할 수 있을 것이다.

본 연구에 한계점은 최근 사회적, 환경적 요인의 변화로 인하여 다양한 형태의 사용량 패턴이 나타날 것이라 예상되었지만, 이에 대한 추후 실측 연구가 아직 제시되어 있지 않은 상황이다. 그러므로, 추후 연구에서는 더욱 정확한 통계적 근거에 입각하여 전체 모집단의 대표성을 가지는 표본추출이 필요할 것이며, 최근 1인 가구 증가 등과 같은 실태를 반영한 독립변수들을 고려하여 이에 대한 용도별 용수자료에 대한 보다 정교한 실측연구가 필요할 것이라 예상된다.

REFERENCES

- [1] Ministry of Environment, "2016 Statistics of Waterworks", pp. 23-24, Ministry of Environment, the Republic of Korea, 2017.
- [2] R. Cochran, A. W. Cotton, "Municipal water demand study, Oklahoma City and Tulsa, Oklahoma," *Water Resources Research*, Vol. 21, No. 7, pp. 941-943, July 1985.
- [3] AWWA Research Foundation, "Commercial and institutional end uses of water", pp. 30-58, American Water Works Association Research Foundation, Denver, 2000.
- [4] Aquacraft, Inc., "Post-drought changes in residential water use", pp. 1-325, Denver Water, Denver, CO., 2006.
- [5] S. L. Zhou, T. A. McMahon, A. Walton, J. Lewis, "Forecasting daily water demand: A case study of Melbourne", *Journal of Hydrology*, Vol. 236, pp. 153-164, 2000.
- [6] L. Shvarster, U. Shamir, M. Feldman, "Forecasting hourly water demands by pattern recognition approach", *Journal of Water Resources Planning and Management*, Vol.

119, No. 6, pp. 611-627, 1993.

[7] J. Adamowski, C. Karapataki, "Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms", *Journal of Hydrologic Engineering*, Vol. 15, No. 10, pp. 729-743, Oct 2010.

[8] M. Nasser, A. Moeini, M. Tabesh, "Forecasting monthly urban water demand using extended Kalman filter and genetic programming", *Expert System with Applications*, Vol. 38, No. 6, pp. 7387-7395, June 2011.

[9] B. Brentan, E. Luvizotto, M. Herrera, J. Izquierdo, R. Perez-Garcia, "Hybrid regression model for near real-time urban water demand forecasting", *Journal of Computational and Applied Mathematics*, Vol. 309, No. 1, pp. 532-541, Jan 2017.

[10] I. Duerr, H. R. Merrill, C. Wang, R. Bai, M. Boyer, M. D. Dukes, N. Bliznyuk, "Forecasting urban household water demand with statistical and machine learning methods using large space-time data: A comparative study", *Environmental Modelling and Software*, Vol. 102, pp. 29-38, April 2018.

[11] J. Kim, H. S. Kim, D. Lee, G. Kim, "Analysis of water use characteristics by household demand monitoring", *Journal of Korean Society of Environmental Engineering*, Vol. 29, pp. 864-869, 2007.

[12] D. Lee, J. Kim, H. Kim, "Statistical analysis of non-household unit water use for business categories", *Journal of Korean Society of Environmental Engineering*, Vol. 29, pp. 385-396, 2009.

[13] H. Kim, J. Kim, "Forecast of regional water demand in Korea: a recursive multi-region water CGE modelling approach", *Journal of the KRS4*, Vol. 26, No. 1, pp. 85-101, Mar 2010.

[14] J. Kim, C. Jung, J. Lee, S. Kim, "Development of naive-bayes classification and multiple linear regression model to predict agricultural reservoir storage rate based on weather forecast data", *J. Korea Water Resour. Assoc.*, Vol. 51, No. 10, pp. 839-852, Aug 2018.

[15] J. Jo, J. W. Baek, "Experimental design and analysis for reliability improvement", *Journal of the Applied Reliability*, Vol. 2, pp. 47-61, 2002.

저자약력

명 성 민(Sungmin Myoung)

[일반회원]



- 2002년 2월 : 연세대학교 일반대학원 의학전산통계학과 (보건학석사)
- 2006년 2월 : 연세대학교 일반대학원 의학전산통계학과 (이학박사)
- 2009년 3월 ~ 현재 : 중원대학교 보건행정학과 부교수

<관심분야> 빅데이터, 데이터 마이닝, 의료정보