

칼만필터의 적응형모델 기법을 이용한 광역상수도 시스템의 수요예측 모델 개발

(The Development of Model for the Prediction of Water Demand using Kalman Filter
Adaptation Model in Large Distribution System)

한태환* · 남익석**

(Tae-Hwan Han · Eui-Seok Nahm)

요 약

본 논문에서는 광역상수도 시스템의 취·송수 설비의 최적운영계획에 필수적으로 요구되는 시간 단위 용수 수요량 예측을 위하여 칼만 필터에 의한 수요 예측 모델 구축 및 배수패턴 해석 기법을 제안하고, 기존 시스템의 실 데이터를 이용하여 시뮬레이션 수행 결과 제안된 기법의 유용성이 검증되었다. 광역상수도 시스템에서 취·송수 설비의 최적운영계획 수립을 위해서는 예측 시간 범위를 최소 하루 단위 이상으로 유지해야 한다. 따라서, 제안된 기법에서는 기존의 시간별 실적데이터의 시계열에 의한 예측을 이용하는 것이 아니라 모델로부터 예측된 일 수요량에 배수패턴을 곱하여 24시간의 시간별 용수 수요량을 예측한다. 일 수요량 예측을 위한 칼만 필터 모델은 입력변수의 통계적 분석에 의해 모델 구조 최적화가 효과적으로 구현되고 배수패턴은 데이터 Granulization에 의해 얻어진다.

Abstract

Kalman Filter model of demand for residential water and consumption pattern were tested for their ability to explain the hourly residential demand for water in metro-politan distribution system. The daily residential demand can be obtained from Kalman Filter model which is optimized by statistical analysis of input variables. The hourly residential demand for water is calculated from the daily residential demand and consumption pattern. The consumption pattern which has 24 time rates is characterized by data granulization in accordance with season kind, weather and holiday.

The proposed approach is applied to water distribution system of metropolitan areas in Korea and its effectiveness is checked.

Key words : Kalman Filter, Residential water, Metro-politan distribution system.

1. 서 론

오늘날 급속한 산업발달과 더불어 도시 인구의 증

가, 도시화의 확대, 생활수준의 향상, 공업 입지의 확대 등으로 용수수요는 계속 증가하는 반면에 수원의 고갈과 수질 오염 등의 확대 등으로 인하여 기존의 수원을 안정된 취수원으로 운용하기가 어려워졌다.

또한, 최근 2~3년간에 경험한 가뭄은 농업용수 뿐 만 아니라 일상생활에 가장 중요한 요소가 되는

* 정회원 : 명지전문대학 전기과 조교수

** 정회원 : LG산전 선임연구원

접수일자 : 2000년 11월 13일

생활용수의 공급에 상당한 어려움을 주게 되어 격일 제급수는 물론, 급수차를 동원한 급수 등 비상수단에 의한 생활용수 공급의 사례가 빈번하게 되었으며, 이를 위하여 광역상수도 체계가 점차 확대되고 있으나, 이에 대한 통합적 계획이나 운영이 이루어지지 않을 경우 과도한 시설 낭비 뿐 만 아니라 소비자인 지방 자치단체간 또는 대규모 공장간에 용수공급상에 공평성 유지에도 문제점으로 제기되고 있다. 따라서, 이러한 용수부족상태에 대한 해결책은 장래 용수수요 예측에 의한 적정 규모의 추가 상수도 시설의 조기 신설과 정확한 용수수요 예측에 의한 공급량을 지방 자치단체별 및 시간대별로 적절히 조절하는 체제가 시급히 요구된다[1].

본 논문은 광역상수도 시스템의 취·송수 설비의 최적운영계획에 필수적으로 요구되는 시간 단위 용수 수요량 예측을 위하여 칼만 필터에 의한 수요 예측 모델 구축 및 배수패턴 해석 기법을 제안한다. 취·송수 설비의 최적 운영계획은 원활한 용수공급이 가능한 범위에서 전력비가 싼 심야시간에 설비(펌프운전)를 많이 운전하여 용수 생산을 많이 하고 전력비가 비싼 시간대에는 설비 운영을 최소화 함으로써 전력비가 절감되게 하는 것이 핵심이다. 따라서 이를 위해서는 실 수용가에의 배수량 예측이 핵심이 되며, 시간별 용수 수요예측을 기존의 시간별 실적데이터의 시계열에 의한 예측을 이용하는 것이 아니라 일 예측량에 정해진 배수패턴을 곱하여 시간별 용수 수요량을 예측하는 방식에 관한 것이다[2][3][4][5]. 즉, 일 단위의 수요량은 기존의 일 단위 실적데이터와 일상생활과 관련된 요일, 기온, 날씨 인자들을 추가하여 시계열 모델링에 의해 예측하고, 이로부터 시간별 수요량은 일 수요량에 해당 배수패턴을 곱하여 시간별 수요량을 예측하는 것으로 효과적인 시간별 수요량이 예측되기 위해서는 최적의 배수패턴의 구현이 핵심기술이 된다. 즉, 해당일의 배수패턴 설정 기준이 예측성능의 핵심요소가 된다. 시간단위 용수 수요예측은 미국과 일본(동경도 수운영 시스템)을 비롯한 대부분의 경우 사용되는 방식이다.

또한, 제안된 기법은 시간별 용수 수요예측 뿐만 아니라 일상생활과 관련된 모든 예측 분야(전력, 전화요금 등등)에 적용이 가능하다. 일반적으로 기존의 예측은 목적하는 시간 단위(년도별, 계절별, 월별, 일별 등등)로 실적데이터에만 의존한 시계열 모델링을

하는 것이 일반적이다. 하지만 이러한 방법은 사회경제적 요인인 인구증가, 실질소득 증가, 주거형태 등의 점진적인 변화들에 의해 영향을 받기 때문에 일상생활과 관련된 시간 단위 또는 생활패턴과 관련된 일정 주기단위의 예측에는 비효율적이다. 또한 이러한 시계열 모델링을 위해서는 예측 인자의 상관관계 분석 및 수치화 작업이 이루어져야 하는 단점이 있다. 따라서, 일상생활과 관련된 분야의 예측은 예측과 패턴을 같이 이용하는 것이 매우 효과적이다. 즉, 예측 시간 단위 보다 이전의 시간단위는 일반적인 예측 기법을 따라서 실적데이터와 환경 요인인 요일, 기온, 날씨 인자를 추가하여 예측을 수행하고, 이 예측량에 일정패턴을 곱하여 시간 단위의 예측량을 산출하는 것이 효과적이다. 특히 용수공급에 있어서는 원활한 용수의 공급과 경제적인 급수시스템의 운영을 위해 용수 수요 단위의 시간별 수요량의 예측이 핵심 기술이 된다. 시간단위로 산출된 용수 수요량은 상위 급수시스템의 운영계획수립 및 펌프운전계획수립의 기반자료가 된다.

그리고, 기존의 용수 수요 예측은 장기예측에 국한되어 이루어졌다. 즉, 년 단위, 월 단위 또는 일 단위의 예측이 주를 이루었고, 시간 단위의 예측은 수행되지 않고 단지 실제 용수를 공급하는 말단 배수지 수위와 연계하여 운전자에 의한 펌프 운전제어에 의해 용수를 공급 받는 형태가 일반적이다. 이러한 경우에는 상위 공급장(취수장, 가압장, 정수장 등등)의 취송수량 계획 및 펌프운전계획이 자동운전이 아닌 수동운전으로 이루어지는 단점과 비효율적인 펌프 운전이라는 단점이 있다. 따라서, 광역상수도의 수운영을 위해서는 시간단위의 수요량 예측이 운영이 기본이 되고 있으며, 일본 동경의 수운영시스템이나 유럽의 수운영시스템은 모두 시간단위의 예측을 기본으로 하여 용수조절 시스템이다.

제안된 기법의 현장시험은 수도권 광역상수도 시스템의 기존 3,4단계 실적 데이터를 이용하여 예측차수결정, 하루 용수 수요량 예측 및 배수패턴 해석을 수행하였으며, SN정수장의 배수지운영에서 실시되었고, 칼만 필터(Kalman Filter)의 모델은 과거 9개월간 정수장에서 용수 생산량과 기상정보로부터 입수한 날씨와 온도를 데이터로 이용하였으며, 우수한 성능을 보여 주고 있다.

2. 광역상수도 시스템

광역상수도시스템의 수요예측은 각 수요대상별 운용 배수지 수위, 자체 생산량, 분기점 유량 등에 의거하여 수요 예측 대상(분기점, 정수장, 배수지)별 일별 수요예측을 실시하고 이를 토대로 각 수요 예측 대상별 배수패턴 해석에 의한 시간별 수요예측량을 산출한다. 수요예측에 따라 산출된 취·송수량은 펌프의 대수에 따라 취·송수량을 수정하고, 수정된 데이터에 따라 각 분기점별 송수량을 산출하여 1일분의 운영계획을 수립 전송하고 전송된 취·송수량에 따라 최적펌프운전 및 밸브제어를 실시하는 부분과, 배수지 및 정수지의 수위는 운영계획에 의해 설정된 수위를 목표치로 하여 운영 제어가 진행되고, 수요예측 오차, 사고 등에 기인하여 배수지의 수위가 운영 폭을 벗어나는 경우에 대하여 이를 보상하는 배수지

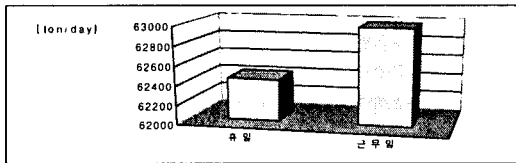


그림 1. SN정수장의 휴일과 근무일의 용수 생산량.
Fig. 1. Water product of work day and holiday in SN water treatment plant.

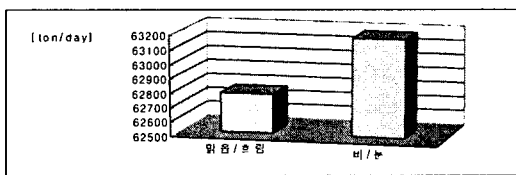


그림 2. SN 정수장의 날씨에 따른 용수 생산량.
Fig. 2. Water product by weather in SN water treatment plant.

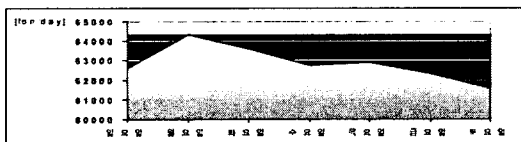


그림 3. SN 정수장의 요일별 용수 생산량.
Fig. 3. Water product according to each day of the week in SN water treatment plant.

수위 안정화 부분으로 구성되어 있다(그림9 참조).

또한, 광역상수도시스템의 배수패턴과 수요량의 변화를 살펴보기 위해 SN정수장의 하루 물 생산 공급량을 날씨별, 근무일과 휴일 및 요일별로 각각 그림1과 그림2 및 그림3에서 보여 주고 있으며, 이 데이터는 9개월간의 평균 1일 데이터를 나타내고 있다.

본 논문에서는 수요예측모델 구축을 위해서는 칼만 필터(Kalman Filter) 기법이 검토되고 배수패턴 해석을 위해서는 데이터 Granulization 기법이 활용된다[6][7][8][9][10][11]. 수요예측모델은 1) 과거 수일간의 실적 배수량, 2) 요일, 3) 당일의 예상 최고 기온, 4) 날씨를 고려한 일별 수요예측량을 산출하고, 이를 기반으로 하여 배수패턴을 이용하여 시간별 수요 예측량을 구한다.

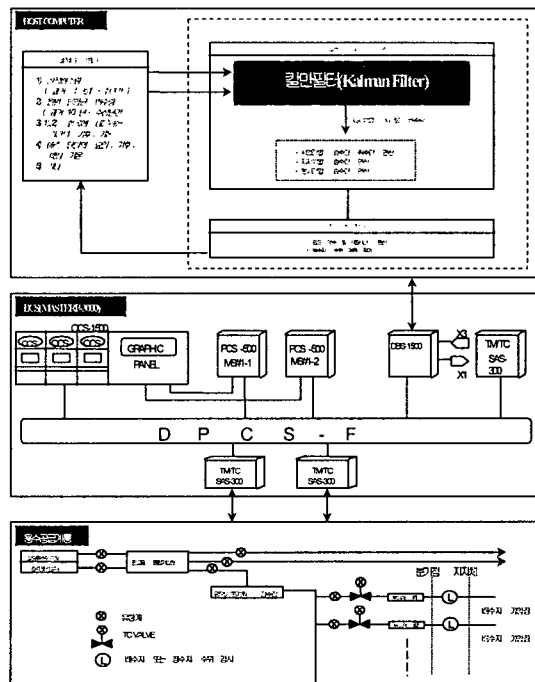


그림 4. 광역상수도 시스템의 전체 구성도.
Fig. 4. Configuration in large distribution system

그림4는 수요예측과 취·송수 제어시스템의 전체적 구성도이다. 중앙 컴퓨터(Host Computer)의 수요예측에 따라 산출된 취·송수량은 펌프의 대수에 따라 취·송수량을 제어하고, 수정된 데이터에 따라 각 분기점별 송수량을 산출하여 1일분의 운영계획을 수

립 전송하며, 전송된 취·송수량에 따라 최적 펌프 운전 및 밸브 제어를 실시하게된다. 또한, 배수지 및 정수지의 수위는 운영계획에 의해 설정된 수위를 목표치로 하여 운영 제어가 진행되고, 수요예측 오차, 사고 등에 기인하여 배수지의 수위가 운영폭을 벗어나는 경우에 대해 이를 보상하는 배수지 수위 안정화 부분으로 구성된다. 용수공급계통에서 유량 관련 데이터(분기점별 밸브 및 펌프운전, 취·송수량 등)가 원격 감시제어(Telemetering/Telecontrol : TM/TC)와 데이터 프로세스 컨트롤러(Data process Control : DPC) 설비를 통하여 분산제어 시스템(Distributed Control System : DCS)으로 전송된다. 전송된 데이터들은 DCS 내의 데이터베이스에 의해 엔지니어링 되어 저장된다. 중앙컴퓨터에서는 이러한 데이터를 이용하여 배수구별 익일 배수량, 시간 배수량, 각 지방단체자치체별 배수량, 분기등을 산출하고, 이를 기반으로 배수구별 취·송수량이 산출된다. 산출 결과는 DCS를 통하여 현장으로 전송된다. 이후의 제어결과가 실적치와 예측치와의 오차가 허용오차 이상인 경우는, 중앙컴퓨터에서 재 연산 작업이 수행되어 산출결과가 재 전송된다.

3. 하루 수요량예측 모델

배수지로부터의 배수량의 시간변동이 예측되면, 배수지의 송수량을 평활화 하거나 배수펌프의 대수 변경 횟수를 최소화하는 송수계획을 세울 수 가있다. 개개의 배수지를 대상으로 각각을 최적화 하는 방법과 여러 개의 정수지를 포함하는 송수계통을 대상으로 최적화 하는 방법을 조합하여 전 도시의 송수유량의 시간계획을 결정한다. 또한, 정수장의 정수지 용량을 활용하는 경우에는 배수지와 마찬가지로 하여 정수량과 취수량의 시간계획을 결정한다. 이들의 계획은 매일 1회의 정시에 자동적으로 동작되도록 하는데, 필요에 따라 매 정시마다 수정하여 이 계획에 따라 동작하도록 한다[12].

본 논문에서의 수요예측모델은 칼만 필터의 선형 예측모델기법이 이용되는데, 데이터베이스를 기반으로 하여 과거의 실적데이터가 이용된 모델이 구축된다. 칼만 필터 예측모델에서는 i 일 이전의 실적배수량 $Q(k-i)$, 요일 $\theta(k)$, 당일(k일)의 예상 최고 기

온 $T(k)$, 당일(k일)의 날씨 $W(k)$ 을 입력변수로 하고, 당일(k일)의 배수량 $Q(k)$ 를 출력변수로 하는 자동회귀이동평균(Auto Regressive Moving Average : ARMA)과 같은 모델식을 구현하고, 이 모델식을 칼만필터 상태 방정식과 출력방정식으로 변형하여 최적의 당일(k일) 배수량 $Q(k)$ 을 산출한다.

수요예측모델 구축을 위한 전체적인 구성은 그림 5와 같다. 데이터 베이스로부터 실적데이터(실적배수량, 요일, 기온, 날씨 등)를 이용하여 모델을 구축한다. 구축된 모델은 필드데이터를 이용하여 월 총수요량, 일 총수요량, 시간별 수요량을 산출하여 취·송수 제어를 위해 데이터베이스에 재 저장한다.

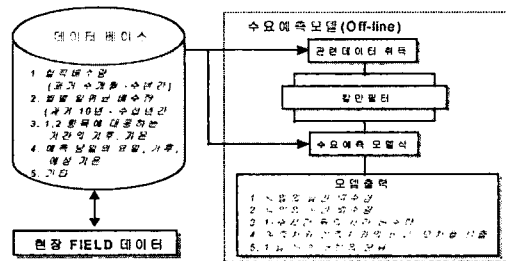


그림 5. 수요예측 모델 구성도
Fig. 5. Block diagram of model for residential water demand

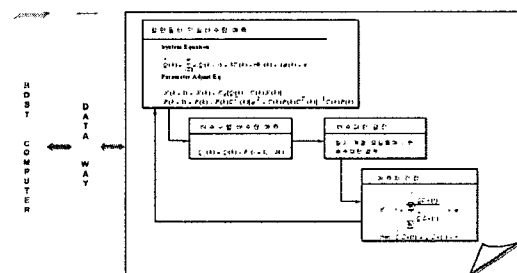


그림 6. 칼만필터법에 의한 배수량 예측시스템
Fig. 6. Kalman Filter for residential water demand

일반적으로 예측기법은 위너(Wiener) 필터계열과 칼만 필터계열 두 가지로 분류된다. 위너 필터계열에는 자동회귀(Auto Regressive : AR)모델, 이동평균(Moving Average : MA)된 모델, ARMA 모델등의 시계열(Time series) 함수들이 포함되고 이들은 과거 데이터들에 대한 상관관계에 의해 현재의 값을 예측

할 수 있는 기법이다[13][14]. 과거의 데이터를 수집하고 최적화된 차수를 결정하여 각각의 과거 데이터에 적절한 가중치를 주어 추정된 현재값을 추정할 수 있다. 칼만 필터는 보다 적응성이 뛰어난 방법으로 과거의 입력값이 현재의 추정값을 결정하기 위해 어떻게 조정되어야 하는지에 대한 정보를 제공한다 [15][16][17].

이러한 과정을 통해 얻어진 결과가 최적화된 추정치가 된다. 칼만 필터는 상태공간기법을 써서 최소자승법을 공식화하는 또 다른 방법을 제시한다. 칼만 필터의 가장 큰 특징은 고려 대상인 랜덤프로세스를 백터화 할 수 있고 측정시 발생하는 오차들을 반복적으로 회귀분석하여 처리할 수 있다는 것이다. 본 연구에서는 수요예측모델로 칼만 필터법을 이용하여 1일량 예측이 처리됨을 보여준다. 그림6은 칼만 필터법에 의한 배수량 예측시스템의 구성도이며, 선형 예측모델에 대한 칼만 필터의 구성식은 다음의 식(1)과 같은데, 신호 x 의 현재값이, 과거의 신호 x 와 현재와 과거의 입력값 u 의 선형결합으로 생각한다. 이와 같은 모델을 선형 예측모델이라고 부른다.

$$x_k = - \sum_{l=1}^n a_l x_{k-l} + G \sum_{l=0}^r b_l u_{k-l} \quad (1)$$

여기서, $a_l=0$ ($1 \leq l \leq n$)일 때 MA 모델이며, $b_l=0$ ($1 \leq l \leq r$)일 때의 모델을 AR 모델이라고 부른다. 양자 모두가 1차이상의 다항식일 때는 ARMA 모델이라고 한다.

일반적으로 단일 입출력의 정 계수 선형의 다이내믹 시스템은 어떤 조건하에서의 AR 또는 ARMA 모델로 나타낸다. 모델의 계수는 조금씩 변화하든지 또는 측정 불가능이든지 하는 것이 일반적이다. 이와 같은 상황에서는 측정데이터를 얻을 때마다 계수의 추적량을 보다 양호한 값으로 변화시켜가는 접근방법이 효과적이고 이런 방법을 적응형 방식이라고 부른다. 칼만필터를 응용하여 ARMA 모델의 계수를 추정하는 적응형의 방식을 유도할 수 가 있다. 그러므로 식(1)의 선형예측모델을 다음의 식(2)와 같이 나타낼 수 가있다.

$$x_k = - \sum_{l=1}^n a_l x_{k-l} + G \sum_{l=0}^r b_l u_{k-l} = \phi f_k \quad (2)$$

여기서,

$$\phi = [- a_1, - a_2, \dots, - a_n, b_0, b_1, \dots, b_r,] \quad (3)$$

$$f_k = [x_{k-1}, - x_{k-2}, \dots, x_{k-n}, u_k, u_{k-1}, \dots, u_{k-r},] \quad (4)$$

실제에 있어서의 측정값 x 는 잡음 w 를 수반하므로 예측모델은 다음의 식(5)와 같이 표현된다.

$$x_k = f_k \phi + w_k \quad (5)$$

여기서, 잡음 w 는 평균값 0 으로의 분산값이 σ^2 이라고 가정하면 ϕ 의 사후 확률밀도의 평균치와 공 분산 행렬은 다음 식(6)과 식(7)과 같이 구해진다.

$$\phi_k = P_k (P_{k-1}^{-1} \phi_{k-1} + \sigma^{-2} x_k f_k) \quad (6)$$

$$P_k = (P_{k-1}^{-1} + \sigma^{-2} f_k f_k')^{-1} \quad (7)$$

여기서, 식(6)과 식(7)을 다르게 표현하면 최종적으로 ARMA모델의 계수 ϕ 의 추정치를 구하여 적응형의 필터인 다음의 식(8)과 식(9)로 나타낼 수 있다.

$$\phi_k = \phi_{k-1} + \sigma^{-2} (x_k + f_k' \phi_{k-1}) P_k f_k \quad (8)$$

$$P_k = P_{k-1} - (\sigma^{-2} + f_k' P_{k-1} f_k)^{-1} T_k \quad \text{where } T_k = P_{k-1}^{-1} f_k f_k' P_{k-1} \quad (9)$$

따라서, 배수지의 배수량 예측 모델을 식(10)과 같이 정의하면 식(3)과 식(4)는 식(11)과 식(12)로 정의된다.

$$\hat{Q}(k) = \sum_{i=1}^m a_i Q(k-i) + bT(k) + cW(k) + d\theta_k + e \quad (10)$$

$$\phi = [a_1 \cdots a_m, b, c, d, e] \quad (11)$$

$$f_x = [Q(k-1), \dots, Q(k-m), T_k, W_k, \theta_k, 1] \quad (12)$$

여기서, $\hat{Q}(k)$ 는 현재 k 에서의 예측 수요량, $Q(k-i)$ 는 i 일 이전의 실적배수량, $T(k)$ 는 당일 최고 기온, $W(k)$ 는 당일 날씨, $\theta(k)$ 는 요일 계수, m 은 예측 차수를 나타낸다.

칼만 필터는 상관계수 a_i, b, c, d, e 를 예측오차에 따라 각각 변화시키고, 식 (8), (9)에 의해 식 (10)을 보정한다(적응형). 예측오차란 중간단계 ($1 < i < m$)에서 예측배수량과 실적배수량과의 차이를 의미한다.

1일 단위의 배수량(일일 배수량)이 예측되면 시간 단위 배수량은 식 (13)에 의해 계산된다.

$$\hat{Q}_i(k) = \hat{Q}(k) \times R_i \quad (13)$$

여기서, $\hat{Q}(k)$ 는 일배수량 예측치, $\hat{Q}_i(k)$ ($i=1, \dots, 24$)는 배수지의 시간 단위 배수량 예측치, R_i 는 일일 시간 단위 배수량 비율을 나타낸다. R_i 는 수년간의 배수구별 실적데이터의 Granulization을 이용한 환산 평균에 의해 결정된 배수패턴이고, 배수패턴은 데이터 Granulization 기법에 의해 계산된다.

4. 배수패턴 해석

시간별 용수 수요예측은 기존의 시간별 실적데이터의 시계열에 의한 예측을 이용하는 것이 아니라 일 예측량에 정해진 배수패턴을 곱하여 시간별 용수 수요량을 예측하는 방식에 관한 것이다. 즉, 일 단위의 수요량은 기존의 일 단위 실적데이터와 일상생활과 관련된 요일, 기온, 날씨 인자들을 추가하여 시계열 모델링에 의해 예측하고, 이로부터 시간별 수요량은 일 수요량에 해당 하는 배수패턴을 곱하여 시

간별 수요량을 예측하는 것으로써 효과적인 시간별 수요량이 예측되기 위해서는 최적의 배수패턴의 구현이 핵심기술이다. 즉, 해당일의 배수패턴 설정 기준이 예측 성능의 핵심요소가 되며, 본 논문에서 배수패턴은 데이터 Granulization 기법에 의해 계산된다.

데이터 수집의 가능성 및 관련 시간과 비용의 절감을 위해 적정수의 샘플 데이터로부터 기대수준 이상의 정확도를 유지하는 예측치 제공을 위해 데이터 Granulization 기법을 이용하여 데이터를 분류하고, 분류된 데이터의 환산 평균에 의해 배수패턴을 설정한다.

데이터 Granulization은 기본적으로 시간계수 및 기후계수를 기초하여 이루어진다. 시간계수는 1) 계절 계수(Season parameter), 2) 일 계수(Day parameter)로 구분되고, 기후계수는 날씨 계수(Weather parameter)로 구분된다. 계절 계수의 s 는 봄/가을, 여름, 겨울로 구분된다. 식 (14)는 계절계수를 정량화한 것이다. 일계수 x 는 근무일과 휴일로 구분된다. 식 (15)은 일계수를 정량화한 것이다. 날씨 계수 w 는 맑음/흐림, 비/눈 2가지로 구분되고 식 (16)과 같이 정의한다.

$$s = \begin{cases} 1 & \text{봄/가을} \\ 2 & \text{여름} \\ 3 & \text{겨울} \end{cases} \quad (14)$$

$$x = \begin{cases} 1 & \text{근무일} \\ 2 & \text{휴일} \end{cases} \quad (15)$$

$$w = \begin{cases} 1 & \text{맑음/흐림} \\ 2 & \text{비 또는 눈} \end{cases} \quad (16)$$

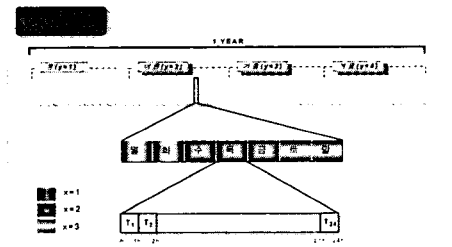


그림 7. 용수 수요예측에서의 계절, 요일 및 날씨에 따른 클래스 분류.
Fig. 7. Class diagram by season, day, weather in water demand prediction.

데이터 Granulization은 시변데이터들을 특성에 따라서 Class로 분류하는 기법이다. 즉, 계절, 요일, 시간 등에 따른 여러 형태의 granule을 구성하고, 구성된 granule을 유사한 형태들끼리 분류하여 Class로 정의한다. 이러한 클래스 데이터를 모델 함수의 과거 데이터로 사용함으로써 효과적인 패턴 설정이 가능하다.

모델화 방법은 접근 방법에 있어서 사회경제적 파라미터들과 다른 시변파라미터들 사이의 커다란 차이를 인식하여 서로 다른 개념에서 출발한다. 사회경제적 파라미터들이 데이터의 빈번한 수집이 불가능한 반면에 시변파라미터들은 특성상 잦은 변화로 인해 통계적 데이터의 수집이 용이하다. 1년 주기의 데이터는 총 8760개(24시간×365일)가 될 것이다. 이 중 계절과 요일적 특성에 의해 같은 형태의 데이터들로 분류가 가능하다. 예를들어, 봄이고 평일이며 같은 시간대인 데이터들은 여러개가 있을 것이다. 이런 식으로 8760개의 총 데이터를 갖은 형태의 데이터들로 분류하고 이러한 데이터들은 각각의 상황에서 특정 클래스의 배수 패턴을 추정하는 과거데이터로 사용된다. 이를 위한 필요조건은 아래와 같다.

- 1) 시간별 모든 granule에 대하여 요구되는 과거 데이터 수집

1 년간의 데이터 수집량 :
 $24(\text{시간}) \times 365(\text{일}) = 8760 \text{ 개}$

- 2) Partitioning 또는 Clustering :

s, x, w 각각에 대한 class의 총 수
 $s(3 \text{ cases}) \times x(2 \text{ cases}) \times w(2 \text{ cases}) + 1(\text{특이일})$
 $= 13(\text{classes})$

그림7은 클래스의 분류 도식을 나타낸다. 이러한 알고리즘을 기초로 하여 과거의 샘플 데이터들을 이용하여 클래스별 배수 패턴을 연산한다. 즉, 단일 클래스내의 요소(element)인 모든 granule을 바탕으로 해당 클래스의 배수패턴이 환산되고 평균되어 연산된다. 특정 클래스의 샘플 granule의 index인 윈도우 크기는 식 (17)과 같다. 그림8은 특정 클래스를 나타낸다.

Window 크기

$$N : t_k = \{ t_1, \dots, t_{p-1}, t_p \} \quad (17)$$
 (t_0 : 시작시점, t_p : 현재시점)

특정 지역에서의 이러한 index들을 이용하여 입력 계수 { s, x, w }들에 의해 할당된 모든 클래스에 관한 샘플 데이터를 수집하여 현재 기준시점 t_p 의 배수 패턴 $\hat{M}_I(t_p)$ 가 식 (18)에 의해 배수패턴이 예측된다. 모델의 결정 요소 $\hat{g}(\cdot)$ 는 환산평균 함수가 된다.

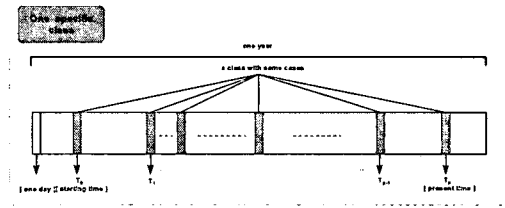


그림 8. 샘플 granule의 윈도우상에서 나타낸 특정 클래스.
 Fig. 8. Specific class in window of sample granule.

$$\hat{M}_I(t_p) = \hat{g}(M_I(t_1), M_I(t_2), \dots, M_I(t_{p-1})) \quad (18)$$

Window 크기는 데이터 granule의 개수가 된다. 이는 데이터의 종류와 수집 방법에 영향을 받으므로 실제적 운영으로부터의 피드백 입력 활용이 가능하다. 또한 정확도와 계산량 간에 상호보완이 고려되어야 한다.

이 경우의 사용자 입력은 1)계절(s)의 선택, 2) 요일 및 날씨 (x, w)이다. 이러한 입력을 바탕으로 데이터의 granulization에 의한 클래스가 결정된다. 예를 들어, 어느해 여름의 비오는 날 주말의 배수패턴의 경우는 다음과 같다. 이 경우 식 (18)에 의한 배수패턴은 아래 3단계 절차에 의해 구해진다.

- 【절차 1】 주어진 지역의 클래스 데이터 집합을 탐색한다.
- 【절차 2】 클래스로 분류된 데이터를 이용하여 구 시간별 샘플데이터를 추출한다.
- 【절차 3】 이러한 샘플데이터를 이용하여 배수패턴을 구한다.

[배수 패턴을 구하기 위한 예제 : 비오는 날 여

름, 주말의 배수 패턴]

【절차 1】 분할된 클래스들의 탐색

$$\{s=2, x=2, y=2\} = M(t_p)$$

여기서, $M(t_p)$: 클래스 set

【절차 2】 클래스 대한 샘플데이터 수집

클래스 :

$$M(t_p) = \{ M_1(t_1), M_1(t_2), \dots, M_1(t_{p-1}) \}$$

【절차 3】 각 클래스의 추정

$$\hat{M}(t_p) = \hat{g}(M_1(t_1), M_1(t_2), \dots, M_1(t_{p-1}))$$

여기서, $\hat{g}(\cdot)$ 는 환산평균 함수

시간별 배수량은 당일의 일기와 계절 및 요일을 조건으로 한 데이터 granulization 기법에 의해 배수 패턴이 결정된다. 배수량 예측치의 검증은 배수구마다 검정 시각까지의 시간과 배수량 예측치 및 실적치에 대한 누계비의 식 (19)로 검증한다.

$$r = \frac{\sum_{n=1}^t QD_n(t)}{\sum_{n=1}^t \hat{Q}D_n(t)} \quad (19)$$

여기서, $QD_n(t)$ 는 시각 t 에서의 n 번째 배수구의 배수량 실적치, $\hat{Q}D_n(t)$ 는 시각 t 에서의 n 번째 배수구의 배수량 예측치를 나타낸다. 식 (19)에 의해 계산된 r 이 허용치보다 크면 예측치의 변경이 필요 없고, 만약 r 이 허용치보다 작으면 예측치의 변경은 (20)을 따른다.

$$\hat{Q}D_n(t) = \hat{Q}D_n(t) \times r \quad (20)$$

5. 시뮬레이션

제안된 칼만 필터법을 이용하여 수도권 광역 상수도 시스템의 기존 3,4단계의 실적 데이터를 이용하여 예측 차수 결정, 1일량 예측 및 배수패턴 해석을 수행하였다. 대상시스템은 한강 수계의 3,4단계의 SN 정수장, SN 배수지, SW 배수지, KH 배수지, WB

정수장, WB 배수지, MG 가압장, GR 배수지를 대상 시스템으로 하여 1일량 예측 및 배수패턴 해석을 수행하였다. 그림 9는 3,4단계의 SN 정수장과 WB 정수장의 계통도를 나타낸다.

수요 예측에서 효과적인 모델 구조의 선택이 매우 중요한 요소이다. 즉, 적절한 입력변수의 선정 및 모델 차수의 결정이 효과적인 수요 예측을 위한 핵심 요소가 된다.

일반적으로 모델 차수 m 은 3에서 7사이의 값을 갖는 것이 일반적이다. 하지만 이의 특별한 결정 방법이 정량화되어 있지는 않고 운전자의 경험에 주로 의존한다. 여기서는 이의 효과적인 구현을 위해 시뮬레이터에 의해 오차를 평가하여 결정하는 방법을 제시한다. 즉, 모델 차수 m 의 값을 증가시켜가며 이에 따른 오차의 변화를 관찰하고 최소의 오차가 되는 모델 차수 m 을 결정한다.

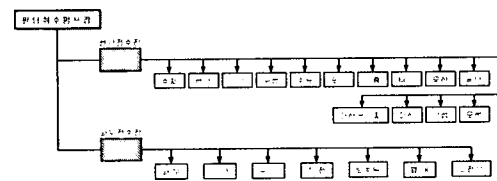


그림 9. 3,4단계 성남 정수장 및 외부 정수장 계통도

Fig. 9. SUNGNAM/WABOO distribution system

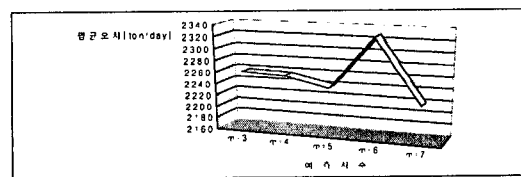


그림 10. 예측차수 (m)의 증가에 따른 평균오차

Fig. 10. Average error with increasing m

따라서 모델 차수 m 의 결정 알고리즘은 다음 순서에 의해 처리된다.

[step 1] 초기값 설정($m=3$)

[step 2] 모델 오차 계산

[step 3] m 의 증가($m = m + 1$)

[step 4] 만약 모델 오차가 기준 설정치 보다 작거나 $m > 7$ 이면 종료.

그렇지 않으면, Go To [Step 2]

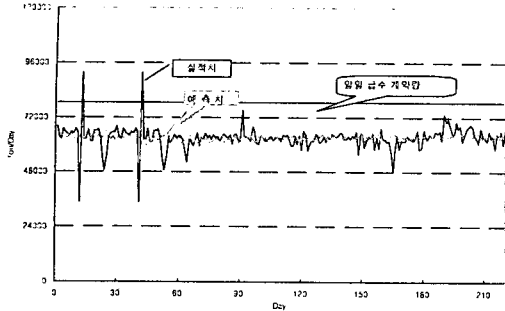


그림 11. SN정수장 유입량 예측치 및 실적치(예측차수=3)
Fig. 11. Output of influent prediction model at SN water purification plant (order=3)

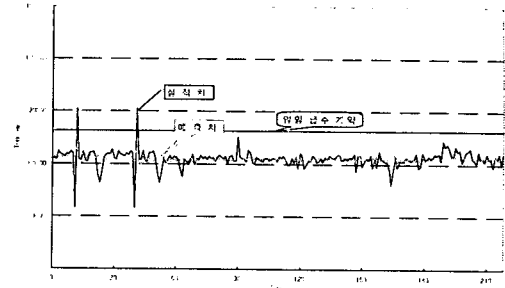


그림 12. SN정수장 유입량 예측치 및 실적치(예측차수=4)
Fig. 12. Output of influent prediction model at SN water purification plant (order=4)

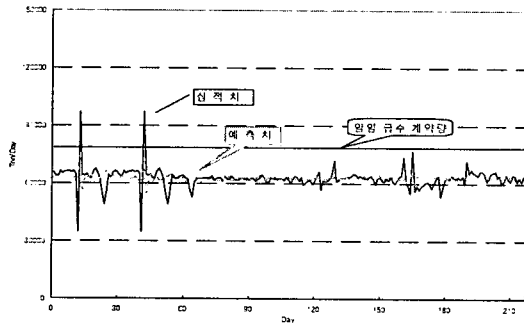


그림 13. SN정수장 배수량 예측치 및 실적치(예측차수=3)
Fig. 13. Output of effluent prediction model at SN water purification plant (order=3)

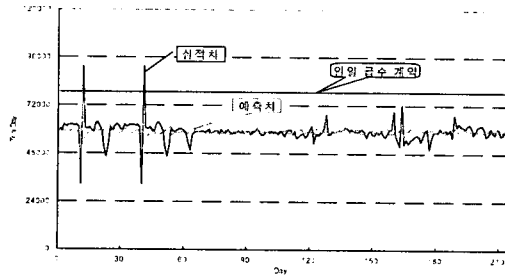


그림 14. SN정수장 배수량 예측치 및 실적치(예측차수=4)
Fig. 14. Output of effluent prediction model at SN water purification plant (order=4)

SN정수장의 유입량 데이터(9개월)를 이용하여 및 칼만 필터법 및 데이터 Granulization기법을 이용하여 예측 차수 m 의 증가에 따른 평균오차를 그림10에 나타내었다. 그림10의 분석 결과 칼만 필터법에 의한 오차변화는 예측차수가 3~5 사이에는 상대적으로 변화가 적다. $m=6$ 에서 오차가 증가한 이유는 일 배수량이 전일 실적치와의 상관관계를 고려할 때 상대적으로 상관정도가 떨어진다는 것을 의미한다. 따라서 1일량 예측에서는 예측 차수를 3과 4에 대해서만 수행하였다. 예측 차수가 5인 경우는 차수 증가에 따른 계산량의 증가를 고려하여 제외하였으며, 입력조건으로는 1) 요일($\theta(k)$), 2) 날씨($W(k)$), 3) 당일 예상 최고 기온($T(k)$), 4) $Q(k-i)$ ($i=3$ or $i=4$)는 1일 전의 실적 유입량 또는 배수량을 입력 조건으로 하여 1일량 예측을 수행하였다.

예측 대상별로 예측 차수 3과 4에 대하여 칼만 필터법을 이용하여 1일량예측을 수행하였다. 시뮬레이션 결과를 대상으로 1) 평균오차, 2) 최대오차, 3) 최소오차, 4) 평균실적치 대비 평균오차의 백분율 4가지로 구분하여 비교평가를 하였다. 그림11, 그림12, 그림13 및 그림14는 SN정수장의 모델값과 실제값의 비교 그래프이다. 모델값과 실제값의 비교 그래프에서 점선은 예측치를 실선은 실적치를 보여주고, 직선은일일 급수 계약(용수의 수요처와 공급처간)량을 나타낸다.

제안된 기법의 시스템별 계약량 및 오차는 표 1~3과 같다. 전반적으로 $\frac{\text{평균오차}}{\text{평균실적치}} \times 100$ 항목을 고려할 때 $10 \pm 5\%$ 내외로 제안된 기법이 실제 시스템에 적용되는 매우 효과적일 수 있을 것으로 생각된다.

표 1. SN정수장 유입량 계약량 및 오차
Table 1. Influent Contraction and Errors of SN Water Purification Plant

오차 항목		계약량	평균오차	최소오차	최대오차	평균실적치	$\frac{\text{평균오차}}{\text{평균실적치}} \times 100$
기법							
칼만 필터 모델	모델차수=3	786,000	29159	189	288932	633,891	4.6 %
	모델차수=4	786,000	28874	120	288801	627,695	4.6 %

표 2. SN정수장 배수량 계약량 및 오차
Table 2. Effluent Contraction and Errors of SN Water Purification Plant

오차 항목		계약량	평균오차	최소오차	최대오차	평균실적치	$\frac{\text{평균오차}}{\text{평균실적치}} \times 100$
기법							
칼만 필터 모델	모델차수=3	786,000	28702	120	282379	623,956	4.6 %
	모델차수=4	786,000	27194	7	265018	632,428	4.3 %

표 3. SN 분기 계약량 및 오차
Table 3. Influent Contraction and Errors of SN Junction

오차 항목		계약량	평균오차	최소오차	최대오차	평균실적치	$\frac{\text{평균오차}}{\text{평균실적치}} \times 100$
기법							
칼만 필터 모델	모델차수=3	110,000	4981	1	47030	113,204	4.4 %
	모델차수=4	110,000	4898	13	44591	111,318	4.4 %

6. 결과고찰

본 논문은 광역상수도시스템의 수요예측을 위해 각 수요대상별 운용 배수지 수위, 용수 생산량, 분기점 유량 등에 의거하여 수요 예측 대상(분기점, 정수장, 배수지)별 1일별 수요예측을 실시하고 이를 토대로 각 수요 예측 대상별 배수패턴 해석에 의한 시간별 수요예측량을 산출한다. 수요예측에 따라 산출된 취·송수량은 펌프의 대수에 따라 취·송수량을 제어하고, 수정된 데이터에 따라 각 분기점별 송수량을 산출하여 1일분의 운영계획을 수립 전송하고, 전송된 취·송수량에 따라 최적펌프운전 및 밸브제어를 실시하는 부분과, 배수지 및 정수지의 수위는 운영계획에 의해 설정된 수위를 목표로 하여 운영 제어가 진행되고, 수요예측 오차 및 사고 등에 기인하여 배수지의 수위가 운영폭을 벗어나는 경우에 대해 이를 보상하는 배수지 수위 안정화를 목적으로 사용된다.

취송수 설비의 최적 운영계획은 원활한 용수공급이 가능한 범위에서 전력비가 싼 심야시간에 설비(펌프 운전)를 많이 운전하여 용수 생산을 많이 하고 전력비가 비싼 시간대에는 설비 운전을 최소화 함으로써 전력비가 절감되게 하는 것이 핵심이다. 따라서, 이의 효과적인 구현을 위해서 수요예측모델 구축에 칼만 필터(Kalman Filter) 기법이 검토되었고 배수패턴 해석을 위해서는 데이터 Granulization 기법이 활용되었다. 수요예측모델은 과거 수일간의 실적 배수량과 요일, 당일의 예상 최고 기온 및 날씨를 고려하여 1일별 수요예측량을 산출하고, 이를 기반으로 하여 배수패턴을 이용한 시간별 수요예측량을 구하였다.

제안된 기법의 실 시스템 적용에 앞서 기존 시스템의 실 데이터를 이용하였으며 시뮬레이션 수행 결과 제안된 기법의 유용성이 검증되었다.

참 고 문 헌

- [1] Ichikawa. A and Ikeda. S, "Prediction of Water Quality by Heuristic Self-Organization" Proceeding of JSCE, No. 246, pp.47~56, 1976.
- [2] Box. G. E and Jenkins. G. m, "Time Series Analysis Forecasting and Control," Holden-day, 1976.
- [3] Melard. G, "Illustration of the Use of a General Time Series Model," TIME SERIES ANALYSIS. THEORY AND PRACTICE 7, pp.53~75, 1985.
- [4] Tsutsumi. M and Chishaki. T, "Time-Series Prediction System and AROP Model in Transportation Demand Analysis," Proceeding of JSCE, No. 407/IV-11, pp17~26, 1989.
- [5] Gray. H. L and Nein. F. Z, "On a Class of Nonstationary Process," JOURNAL OF TIME SERIES ANALYSIS, Vol. 9, No. 2, pp.133~154, 1988.
- [6] C. T. Leondes, (Ed.), "Theory and Application of Kalman Filtering," North Atlantic Treaty Organization AGARD Report, No. 139, February 1970.
- [7] H. W. Sorenson, "Kalman Filtering Techniques," in Advances in Control Systems, Vol. 3, C. T. Leondes (Ed.), New York: Academic Press, pp.219~289, 1966.
- [8] H. W. Sorenson (Ed.), "Kalman Filtering: Theory and Application," New York: IEEE Press, 1985.
- [9] A. A. Girgis, "Application of Kalman Filtering in Computer Relaying of Power Systems," Ph. D. dissertation, Iowa State University, Ames, 1981.
- [10] A. A. Girgis and R. G. Brown, "Application of Kalman Filtering in Computer Relaying," IEEE Trans. on power Apparatus and Systems, PAS-100: No. 7, pp.3387~3397, July 1981.
- [11] Ardeshir Faghri, Sandeep Aneja, "Artificial Neural Network-Based Approach To Modeling Trip Production", Transportation Research Record, Vol. 1556, 131-136.
- [12] Kubota Masakazu , Kurotani Kenichi, Akiyama Hirohide, Morimoto Masanori, "Altitude of Water Operation Control Package", Journal of EICA Vol.1. No.1,1996.
- [13] Anderson. O. D, "Practical Examples of ARIMA Modelling," pp.15-21, 1985. Demand," Proceeding of JSCE (submitted for publication)
- [14] Cholette. P. A, "Seasonal Moving Average for Irregular in the Series and with Moving Seasonality," TIME SERIES ANALYSIS. THEORY AND PRACTICE 7, pp.37~52, 1985.
- [15] A. H. Jazwinski, "Stochastic Processes and Filtering Theory," Academic Press, New York, 1970.
- [16] Robert G. Brown, Patrick Y. C. whang, "Introduction to Random signals and Applied Kalman Filtering", Second Edition, John Wiley & Sons, 1992.
- [17] Jerry M. Mendel, "Lessons in Estimation Theory for Signal Processing, Communication, and Control", Prentice Hall PTR. P121~P131, 1995.

◇ 저자소개 ◇

한 태 환 (韓泰煥)

1956년 3월 18일생. 1980년 한양대학교 전기공학과 졸업. 1986년 동대학원(석사) 졸업. 1998년 충북대학교 대학원(공학박사) 졸업. 1980~1998년 LG산전 근무. 1999년 서울대 계측제어신기술연구센터 근무. 2000년 3월~현재 명지전문대학 전기과 교수. 공업계측제어 기술사.

남 의 석 (南宜錫)

1968 2월 1일생 1991년 연세대 전기공학과 졸업(공학사), 1993년 연세대 전기공학과 대학원 졸업(석사), 1998년 연세대 전기공학과 대학원 졸업(박사), 현재 LG산전 주식회사 선임연구원. 주요관심분야 : AI, System Integration, 수요예측, 지능형 제어, 모델링, Water management System.