



상수관망의 누수감지를 위한 주성분 분석의 적용 가능성에 대한 연구

Study on the applicability of the principal component analysis for detecting leaks in water pipe networks

김기민·박수완*
Kimin Kim·Suwan Park*

부산대학교 공과대학 사회환경시스템공학과, 부산광역시 금정구 부산대학로63번길 2, 46241
*Department of Civil and Environmental Engineering, Pusan National University, 2, Busandaehak-ro 63beon-gil,
Geumjeong-gu, Busan, Republic of Korea, 46241*

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 159-167

ABSTRACT

In this paper the potential of the principal component analysis(PCA) technique for the application of detecting leaks in water pipe networks was evaluated. For this purpose the PCA was conducted to evaluate the relevance of the calculated outliers of a PCA model utilizing the recorded pipe flows and the recorded pipe leak incidents of a case study water distribution system. The PCA technique was enhanced by applying the computational algorithms developed in this study which were designed to extract a partial set of flow data from the original 24 hour flow data so that the effective outlier detection rate was maximized. The relevance of the calculated outliers of a PCA model and the recorded pipe leak incidents was analyzed. The developed algorithm may be applied in determining further leak detection field work for water distribution blocks that have more than 70% of the effective outlier detection rate. However, the analysis suggested that further development on the algorithm is needed to enhance the applicability of the PCA in detecting leaks by considering series of leak reports happening in a relatively short period.

Key words: Principal component analysis, Water distribution blocks, Water pipe network, Leak, Computational algorithm, Flow data

주제어: 주성분분석, 상수도 배수 블록, 상수도 관망, 누수, 계산 알고리즘, 유량 자료

1. 서 론

최근 상수도 관망의 누수, 관로파열, 계량오차로 인한 무수수량을 감소시키기 위한 스마트 물관리(smart water management) 기술 개발이 세계적으로 활발히 진행되고 있다. 스마트 물관리 기술의 한 분야로서, 상수도 관로의 누수량을 산정하는 기법 개발에 관한 연구 사례를 보면 배수 관망의 누수량을 산정하기 위한

top-down 방법에 기반한 IWA법, Covas et al. (2006)의 bottom-up 법을 이용한 야간 최소 유량 분석법이 있다. 또한 상수도 관망의 누수탐지 기법은 Kapelan et al. (2004) 및 Stathis and Loganathan (1999)의 최적화 및 수리학적 모형을 이용한 기법, Muggleton et al. (2006)의 음향 로깅, Pilcher et al. (2007)의 단계적 시험법, Muggleton and Brennan (2005)의 상관관계법이 있다.

상수도 관망의 유량 자료를 이용한 알고리즘을 통하여 상수도 관망의 누수를 판단하는 기법 개발에 관

Received 13 July 2017, revised 26 October 2018, accepted 8 April 2019.

*Corresponding author: Suwan Park (E-mail: swanpark@pusan.ac.kr)

한 연구 사례는 인공지능망을 적용하여 누수를 예측하는 기법(Mounce et al., 2002), Fuzzy 추론 기법과 수리학적 시뮬레이션을 적용한 기법(Xia and Guojin, 2010), 누수 탐지를 위하여 서포트 벡터 머신을 적용한 기법(Borges and Ramirez, 2010; Tajima and Mita, 2008), 유량과 수압을 이용해 칼만 필터링 기법을 적용한 기법(Ye and Fenner, 2011) 등이 있다.

본 연구에서는 다변량 자료분석 통계기법의 하나인 주성분 분석(principal component analysis, PCA)을 상수도 관망에서 관측된 유량자료에 적용하여 검출된 이상치와 상수도 관로 유지관리 기록상의 누수와 관련성을 평가하여, 상수도 관망의 누수 탐지를 위한 주성분 분석의 적용 가능성을 분석하였다.

상수도 관망의 유량관측 자료의 분석에 주성분 분석기법을 적용한 최초의 연구는 Palau et al. (2003)로서, 이들은 상수도 관망에서 관측한 유량 자료를 이용하여 주성분 분석을 실시한 후 T²Hotelling 통계치와 DMOD(distance to model) 통계치를 구하고 이를 각 통계치의 임계값과 비교함으로써 주성분 분석을 이용한 누수감지를 시도하였다.

국내에서는 Park et al. (2013)이 국내 연구대상 실제 관망의 4개 소블록에서 4개월 동안 관측된 야간유량 자료에 대해 주성분 분석을 적용하여 관망의 이상징후를 판정하기 위한 최적의 주성분 분석기간 설정에 관한 연구를 실시하였으며, Park et al. (2015)은 주성분 분석 기법을 활용하여 주로 야간 시간대에 행해지는 누수탐사를 보다 효율적으로 할 수 있게 하는 상수도 관망 블록내의 대표적인 용수사용 유형에 대한 최소 용수사용 시간을 결정하기 위한 컴퓨터 계산 알고리즘을 개발하였다.

이와 같이 상수도 관망의 누수 예측 또는 탐지에 인공지능 알고리즘을 이용한 연구는 비교적 최근에 시도되고 있다. 그러나 칼만필터링의 경우에는 자료가 누락되는 경우 분석이 불가능하며, 인공지능망을 이용한 방법의 경우 상수도 관망의 특성 예를 들어 밸브 개폐, 관망 확장 및 개보수 등이 바뀌게 되는 경우 새로운 관측 자료를 이용하여 인공지능망을 새로이 훈련시켜야 한다. 마지막으로 서포트 벡터 머신의 경우 인공지능을 훈련시키는데 필요한 자료군을 확보하기 위해 실제 상수도 관망에서 수백~수천건의 인공 누수를 발생시켜야 한다. 따라서 위의 방법을 이용한 자동누수감지의 경우에는 실무에 적용하는 것이 실질

적으로 불가능하므로 유량의 결측치가 존재하는 경우에도 분석이 가능한 주성분 분석을 이용한 연구가 필요하다.

본 연구에서는 관망의 이상징후 분석에 주성분 분석기법을 적용한 Palau et al. (2003)의 연구방법론을 확장하여 국내의 실제 상수도 관망에서 관측된 유량 자료를 주성분 분석에 적용하였다. 관측된 유량자료 중 이상치를 판별한 다음 이를 누수기록과 비교함으로써 주성분 분석 모델의 상수도 관망 누수감지 적용 가능성을 분석하였다. 특히 이상치 탐지 유효율을 향상시켜 누수탐지에 주성분 분석 기법의 적용성을 높일 수 있는 계산 알고리즘을 개발하였으며 이를 컴퓨터 프로그래밍하여 연구대상 관망의 유량관측 자료 주성분 분석에 적용하였다. 본 연구와 비슷한 주제로 연구한 Park et al. (2013)은 주성분분석의 이상치를 이용한 관망의 이상감지 유효율의 계산에 유량자료의 시간범위를 한정하지 않고 하루 24시간 유량자료 전체를 사용하였으나, 본 연구에서는 이상치 탐지 유효율을 향상시키기 위하여 각 블록에 대하여 측정된 유량자료 전체를 이용하는 것이 아니라 특정 시간대의 유량자료만을 사용하였다.

2. 연구방법

2.1 주성분 분석의 개념

주성분 분석은 다변량 자료를 해석하는 분석기법 중 하나로서 다차원 자료에 포함된 정보의 손실을 최소화하여 낮은 차원의 자료로 변환하는 기법이다. 원래 자료의 변수의 개수보다 더 적은 개수의 변수를 통하여 전체의 정보를 표현하고자 하는 것이 주성분 분석의 주요 개념이다. 주성분(principal components, PC)들은 서로 통계적인 의미로 독립적이며, 모든 주성분을 사용하여 원자료를 재추정 할 경우 정보의 손실이 없게 된다. 제 1 주성분이 자료의 변동특성을 가장 잘 설명하며, 나머지 주성분은 제 1 주성분으로 설명하지 못한 자료의 변동성을 설명하며, 주성분의 위계가 낮아질수록 원자료에 대한 설명력은 점차 줄어든다.

주성분 분석은 $n \times m$ 차원(여기서, n 은 관찰의 수, m 은 변수의 수를 의미)을 가지는 원변인에 대하여 표준화된 행렬 X 를 주성분 점수(scores) 행렬 $T(n \times f$ 차원)과 부하(loading) 행렬이라고 불리는 $m \times f$ 차원의 행렬 P



및 잔차(residual)행렬 $E(n \times m$ 차원)로 분할(factorization)하는 방법으로 Eq.(1)과 같다.

$$X = T \times P^T + E \quad (1)$$

여기서 f 는 주성분의 개수이며, $f < m$ 의 관계를 갖는다. 부하 행렬 P 의 열(column)은 X 의 분산-공분산 행렬의 고유값(eigenvalue)에 대한 고유벡터(eigenvector)를 나타낸다.

Singular value decomposition이론에 의해 원 유량자료는 Eq.(1)과 같이 부하행렬과 점수행렬, 잔차행렬로 분해되며 부하행렬의 열벡터, 즉 고유벡터에 대응하는 고유값의 크기가 큰 순서대로 제 1주성분, 제 2주성분, 제 3주성분 등으로 정렬된다. 여기서 원자료의 차원을 낮은 차원으로 낮추는데 사용되는 설명력의 기준값(예를 들어 70~80%)에 따라 분석에 사용되는 주성분의 개수를 선택하면 원자료의 차원이 낮은 차원으로 분해된다.

유량자료의 주성분 분석에 따른 이상치를 판정하기 위하여 자료의 분산을 이용하여 비교적 큰 이상치(strong outlier)를 판정하는 T^2 Hotelling 통계치를 이용하는 방법 또는 주성분 분석 모델과 관측 값의 오차를 이용하여 비교적 크지 않은 보통의 이상치(moderate outlier)를 판정하는 DMOD 통계치를 이용할 수 있다. Palau et al. (2003)에 의하면 T^2 Hotelling 통계치를 이용한 이상치는 관망에서 비정상적 수요를 감지하는데 보다 적합하고, DMOD 통계치를 이용한 이상치는 관망에서 발생한 누수를 감지하는데 보다 적합하다. 본 연구에서는 DMOD 통계치를 기준으로 연구대상 관망 유량자료의 이상치를 분석하였다. DMOD 통계치는 $(S_i/S_0)^2$ 로 계산된다. S_i 와 S_0 는 각각 Eq.(2) 및 Eq.(3)와 같다.

$$S_i = \sqrt{\frac{\sum_{k=1}^K e_{ik}^2}{(K-A)}} \quad (2)$$

$$S_0 = \sqrt{\frac{\sum_{i=1}^N \sum_{k=1}^K e_{ik}^2}{(N-A-A_0)(K-A)}} \quad (3)$$

여기서 S_i 는 모델에 의한 추정치의 절대 오차, S_0 는 모델에 의한 추정치의 정규화된 오차, K 는 원자료의

변수의 개수, A 는 주성분의 개수, N 은 관측 자료의 개수를 나타내며, A_0 는 표준화 했다면 1, 그렇지 않다면 0의 값을 가지고, e_{ik} 는 Eq.(1) 잔차행렬 E 의 i 번째 행 k 번째 열의 성분값이다. Eq.(4)은 산정된 DMOD 통계치가 이상치 인지를 판정하는 기준을 나타낸다.

$$(S_i/S_0)^2 > F_{critical} \quad (4)$$

여기서 $(S_i/S_0)^2$ 는 $(N-A-1)(K-A)$ 자유도를 가지며 $(K-A)$ 를 매개변수로 하는 F-Snedecor 확률분포를 가진다. 따라서 계산된 DMOD통계치 $((S_i/S_0)^2)$ 가 F-Snedecor 확률분포값의 유의수준 0.05에 해당하는 값, 즉 $F_{critical}$ ($p=0.05$)보다 크면 i 번째 관측일의 24시간 유량자료전체에 대한 DMOD 통계치는 이상치로 판정된다. 따라서 본 연구에서는 i 번째 관측일의 블록 유입유량을 누수와 같은 관망의 이상징후를 나타내는 유량으로 의심할 수 있는 근거로 Eq.(4)를 사용하였다.

2.2 사용된 유량관측 자료

본 연구는 국내 N도시 상수도 관망의 11개 소블록 유입량 자료를 사용하였다. 본 연구의 상수도 관망의 유량측정 자료는 2011년 9월 1일부터 2012년 10월 30일까지의 426일 동안 매시간 마다 기록된 24시간 유량자료이다. N시 상수관망 누수관리기록에는 상수도 관망의 유량측정 자료뿐만 아니라 누수 발생 및 복구에 관한 내용이 해당 날짜에 대해 기록되어 있다. 본 연구에 사용된 11개 소블록의 누수관리기록에

Table 1. Sample inflow data of a small block in the case study area water pipe network(m^3/hr)

Date	hour	Block A	Block B	Block C
2012-06-19	23:00	63	19	101
2012-06-19	24:00	36	11	23
2012-06-20	1:00	23	12	88
2012-06-20	2:00	23	12	13
2012-06-20	3:00	24	8	12
2012-06-20	4:00	22	9	12
2012-06-20	5:00	31	11	19
2012-06-20	6:00	44	21	35
2012-06-20	7:00	80	94	112
2012-06-20	8:00	68	148	169
2012-06-20	9:00	71	94	165

존재하는 누수 발생 및 복구 관련 기록은 총 120건이다. Table 1은 연구대상 도시의 상수도 블록 유입량 자료의 일부를 나타낸다.

3. 이상치 탐지유효율 분석

3.1 24시간 유량자료를 이용한 PCA

먼저 연구대상 관망의 각 블록의 유입유량자료에 대하여 PCA를 실시하였으며, 각 블록에 대하여 Eq.(1)의 주성분 점수(scores) 행렬 T, 부하(loading) 행렬 P 및 잔차(residual)행렬 E를 구하였다. 각 블록의 누수 관련 기록이 존재하는 날짜와 PCA모델의 DMOD 통계치의 이상치로 판정되는 날짜를 비교하였다.

연구대상 관망의 누수관련 기록이 존재하는 날짜가 실제 누수가 발생한 날짜가 아닐 가능성이 있으므로, 누수로 신고된 날짜의 전, 후 각 7일 내의 날짜들을 PCA의 이상치 탐지 날짜와 비교하였다. 이는 해당 연구지역에서 누수 사고 발생 후 최대 7일이 지난 후에 누수를 인지하거나 누수 복구 사항을 기록할 수 있으며, 누수 사고 또는 고객 불만 사항을 기록하는 데 오류로 인하여 실제 누수 관련 기록이 누수가 발생하기 최대 7일 전에 기록 될 수 있다는 가정에 근거한다. Eq.(5)는 ‘이상치 검출 기간 내의 탐지유효율(effective outlier detection rate, EODR)’을 산정하는데 적용한 식을 나타낸다.

$$EODR(\%) = \frac{\text{No. of Outliers in ODP}}{\text{Total No. of Outliers}} \times 100 \quad (5)$$

여기서 이상치 검출기간(outlier detection period, ODP)은 누수 기록의 전후 7일을 포함한 15일의 기간이다. Table 2는 연구대상 관망 유량자료에 대한 EODR을 분석한 결과를 나타낸다.

Table 2에 보인 바와 같이 24시간 유량자료를 모두 사용한 PCA의 경우에는 대부분의 블록에서 EODR이 60%이하 인 것으로 분석된다. 특히 E, F 블록의 경우 10%정도의 매우 낮은 유효율을 가진다. 따라서 본 연

구에서는 매시간 관측된 24시간 유량자료 전체를 이용하는 것이 아니라 특정 시간대의 유량자료만을 사용하여 PCA의 관망 누수감지 적용성을 높일 수 있는 컴퓨터 계산 알고리즘을 개발하였다.

연구대상 11개 블록에 대해 24시간 유량자료를 사용한 PCA의 제 1주성분을 주성분 요소값에 따라 구분할 경우 주성분 요소값의 편차가 큰 경우와 작은 경우로 나눌 수 있으며, Fig. 1은 이러한 예로 F와 J 블록에 대한 제 1주성분(부하행렬의 열벡터)의 각 요소값을 나타낸 것이다. Fig. 1에서 F블록의 제 1주성분을 분석하면 F블록은 주간과 야간의 사용유량의 편차가 매우 큰 것을 알 수 있다. 따라서 F블록의 제 1주성분의 의미는 ‘주·야간 사용량 편차’로 볼 수 있다. 한편, J블록의 제 1주성분을 분석하면 J블록의 시간대별 사용유량의 편차가 매우 작으므로 J블록의 제 1주성분의 의미는 ‘24시간 총 유량’으로 정의될 수 있다.

본 연구의 주성분 분석 모델을 이용하면 주성분의 의미를 이용하여 블록의 유입유량자료의 순위를 결정할 수 있다. F블록의 경우 일간 유량 자료들을 상대적으로 평가하여 그 순위를 매기고자 할 때 가장 유용한 기준이 주·야간 사용량 편차인 것을 알 수 있다. 즉, F 블록의 일간 유량 자료 중 ‘주·야간 사용량 편차’가 가장 큰 일간 유량자료가 1위가 되고 그 편차가 작은 순서대로 순위가 정해질 수 있다. 한편, J블록의 일간유량 자료들을 상대적으로 평가하여 그 순위를

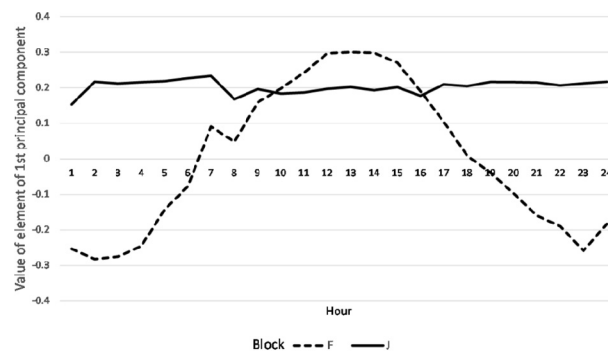


Fig. 1. Values of the first principal component.

Table 2. EODRs using 24hour flow data for the case study water pipe network

Name of block	A	B	C	D	E	F	G	H	I	J	K
No. of leak records	17	12	9	5	9	6	8	2	17	10	25
EODR using 24hour flow data (%)	54	55	51	32	10	9	38	15	19	39	92



정하고자 할 때 가장 유용한 기준은 ‘24시간(일간) 총 유량’으로 볼 수 있다. 즉 J블록의 일간 유량 자료 중 일간 총 유량이 가장 큰 자료가 1위가 되고 그 크기가 작은 순서대로 순위가 정해질 수 있다. 이와 같이 상수도 유량자료의 경우 주성분의 의미는 ‘각 시간대 유량에 대한 가중치’인 것으로 사료된다.

3.2 PCA의 관망 누수감지 적용성을 높일 수 있는 알고리즘

본 연구에서는 각 블록에 대하여 측정된 유량자료 전체를 이용하는 것이 아니라 특정 시간대의 유량자료만을 사용하여 이상치 탐지 유효율을 향상시키는

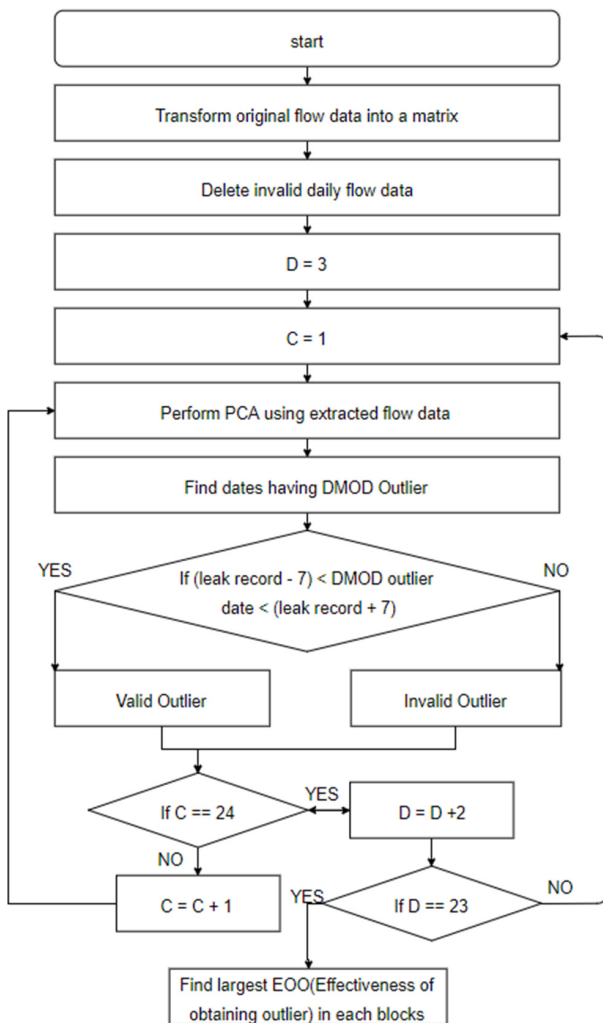


Fig. 18. Computational algorithms for assessing the capability of the PCA for detecting leaks in water pipe network blocks.

방법을 개발하였다. 여기서 시간대는 시간의 범위와 그 범위의 중심 시간으로 정의 된다. 예를 들어 3시부터 7시까지의 시간대인 [3, 4, 5, 6, 7]의 중심시간은 5시, 시간 범위는 5시간이다. 시간대가 13시부터 19시까지인 [13, 14, 15, 16, 17, 18, 19]의 경우는 중심시간은 16시이며 시간범위는 7시간이 된다. Fig. 2는 본 연구에서 개발한 컴퓨터 계산 알고리즘을 나타낸다.

Fig. 2와 같이 본 연구에서는 모든 중심 시간과 시간 범위의 조합에 대해서 누수탐지 유효율을 구하였다. Fig. 1에서 시간 범위(D)는 3시간부터 시작하여 2시간씩 증가하며 23시간까지 커지게 하였으며, 각 시간 범위의 중심 시간(C)을 1시부터 24시까지 변하게 하여 정의되는 각 시간대마다의 유효율을 구하여 각 블록별로 가장 높은 유효율을 갖는 시간대를 탐색하는 알고리즘을 개발하였다. 또한 이상치 탐지 유효율을 일정치 이상으로 갖기 위해 필요한 최소한의 자료를 찾아내기 위해 분석되는 자료의 양에 따른 유효율의 변화 양상에 대해 분석하였고 이를 이용하여 탐지 유효율의 변화와 가장 높은 유효율을 갖는 중심시간, 시간간격의 변화 양상에 대해 분석하였다.

3.3 이상치 탐지 유효율 분석결과

Table 3은 사용된 유량자료가 존재하는 2011년 9월 자료부터 1개월씩 유량자료를 누적하면서 분석하여 2012년 10월까지 과거 자료부터 순차적으로 본 연구에서 개발한 알고리즘을 적용하여 계산한 최고 이상치 탐지 유효율을 나타낸다.

Table 3의 Block B의 month 6에 해당하는 값 ‘0.87’은 B Block의 2011년 9월부터 2012년 2월까지의 유입 유량자료에 대해 중심시간과 시간간격을 변동시켜 구한 이상치 탐지 유효율 중 가장 높은 유효율 87%를 나타낸다. Table 4는 Table 3에 보인 최고 유효율을 가지는 중심시간과 시간간격을 나타낸다.

Table 3과 Table 4를 분석하면 거의 모든 블록이 누적된 유량자료가 부족한 1~4개월의 경우 누수자료가 없거나 그 개수가 적어 탐지유효율, 중심시간, 시간간격 모두 불안정한 양상을 보이나, 가을과 겨울(2011년 10월~2012년 2월)을 지나면서 누수기록이 상대적으로 많이 증가하여 6개월치 자료를 사용한 경우부터는

Table 3. Maximum EODRs for the flow data In chronological order(100%)

Block	month													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	2	3.5	2.45	2.09	2.08	2	1.86	1.7	1.48	1.25	1.32	1.25	1	1
B	1	0.86	0.8	0.89	0.92	0.87	0.88	0.88	0.87	0.73	0.76	0.75	0.73	0.71
C	2	1.4	1.45	1.38	1.39	1.33	1.18	1.2	0.98	1.02	1.04	1.04	1.07	1.04
D	N/A	2.6	2.33	2	1.65	1.27	0.96	0.83	0.78	0.69	0.67	0.65	0.62	0.61
E	4	3	2.14	2.13	2.13	2.13	1.89	1.89	1.7	1.55	1.55	1.42	1.31	1.31
F	N/A	N/A	N/A	0.5	0.5	0.42	0.43	0.47	0.53	0.46	0.5	0.5	0.45	0.45
G	1	1	1.4	0.85	0.83	1	0.94	0.94	0.76	0.56	0.52	0.52	0.54	0.52
H	N/A	N/A	N/A	0.3	0.3	0.27	0.21	0.19	0.17	0.17	0.16	0.16	0.12	0.13
I	2	2	1.8	1.6	1.5	1.3	1.38	1.39	1.33	1.13	1.13	1.07	0.97	0.93
J	2	1.8	1.67	1.47	1.33	1.27	1.07	1	0.94	0.81	0.71	0.68	0.68	0.63
K	2	3	1.78	2.25	2	1.8	1.56	1.5	1.32	1.31	1.32	1.3	1.26	1.24

Table 4. Center time and time duration that correspond to the maximum EODRs

Block		month													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	Center Time	17	5	5	13	15	17	15	15	20	5	5	5	5	5
	Time Duration	3	3	3	11	13	11	15	11	7	3	3	3	3	3
	No. of Leaks	2	7	7	9	12	13	13	14	15	15	16	16	16	17
B	Center Time	14	21	7	8	12	1	12	12	12	11	11	12	13	12
	Time Duration	3	5	11	3	3	3	3	3	3	11	11	7	7	7
	No. of Leaks	1	2	3	5	8	8	8	8	9	9	10	10	10	12
C	Center Time	4	12	22	22	23	10	22	11	22	22	22	22	21	21
	Time Duration	3	7	7	7	7	3	5	5	9	9	9	9	7	7
	No. of Leaks	2	3	6	7	8	8	8	8	8	8	9	9	9	9
D	Center Time	N/A	20	21	20	14	8	11	7	4	4	4	4	4	6
	Time Duration	N/A	5	5	3	9	3	15	3	3	5	5	5	5	3
	No. of Leaks	0	3	3	4	4	4	4	4	4	4	4	4	4	5
E	Center Time	22	22	21	4	5	5	5	5	5	5	5	5	5	5
	Time Duration	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	No. of Leaks	5	6	6	7	7	9	9	9	9	9	9	9	9	9
F	Center Time	N/A	N/A	N/A	18	18	16	18	18	18	10	10	24	23	24
	Time Duration	N/A	N/A	N/A	5	5	3	5	5	5	3	3	3	3	3
	No. of Leaks	0	0	0	1	1	1	1	4	5	5	6	6	6	6
G	Center Time	9	11	21	21	11	20	20	20	20	10	21	21	20	21
	Time Duration	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	No. of Leaks	1	3	4	5	5	6	7	7	7	7	7	7	7	8
H	Center Time	N/A	N/A	N/A	17	15	12	12	12	9	16	19	16	17	16
	Time Duration	N/A	N/A	N/A	3	5	5	5	5	3	7	7	7	5	3
	No. of Leaks	0	0	0	1	1	1	1	1	1	1	2	2	2	2
I	Center Time	4	18	1	1	19	1	2	20	20	19	19	20	1	2
	Time Duration	3	3	5	5	3	3	5	3	3	3	3	3	3	3
	No. of Leaks	3	5	5	8	9	9	11	13	15	15	16	17	17	17
J	Center Time	12	21	1	11	22	9	1	1	1	24	1	1	1	1
	Time Duration	3	5	3	9	7	3	7	7	7	5	7	7	9	7
	No. of Leaks	2	5	7	8	9	9	9	9	9	10	10	10	10	10
K	Center Time	2	17	22	18	17	17	18	17	17	15	18	18	19	19
	Time Duration	3	19	7	21	15	13	21	17	17	3	21	21	17	13
	No. of Leaks	4	6	9	13	14	14	17	17	19	21	22	22	22	25

**Table. 5** Maximum EODRs for the flow data in reverse order of time(100%)

Block	month													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	1.0	0.5	0.6	0.5	0.3	0.4	0.41	0.37	0.4	0.56	0.52	0.54	1.16	1
B	1.0	1.0	0.86	0.86	0.63	0.7	0.63	0.48	0.29	0.67	0.77	0.68	0.76	0.71
C	N/A	N/A	N/A	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.36	1.0	0.83	1.04
D	1.0	0.5	1.0	0.67	0.67	0.4	0.4	0.2	0.25	0.3	0.25	0.21	0.65	0.61
E	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.6	0.6	0.71	0.67	0.56	1.31
F	N/A	N/A	N/A	0.67	0.63	0.67	1.08	0.94	0.9	0.75	0.74	0.65	0.57	0.45
G	1.0	1.0	0.67	0.6	0.5	0.36	0.29	0.33	0.42	0.36	0.38	0.46	0.54	0.52
H	N/A	N/A	N/A	0.33	0.2	0.2	0.11	0.09	0.11	0.08	0.14	0.23	0.18	0.13
I	N/A	N/A	0.17	0.17	0.17	0.4	0.56	0.82	0.68	0.71	0.83	0.82	0.81	0.93
J	N/A	N/A	N/A	N/A	N/A	0.3	0.25	0.2	0.22	0.32	0.28	0.43	0.61	0.63
K	3.0	2.3	2.0	1.86	2.0	1.47	1.10	1.12	1.0	0.93	1.10	1.15	1.17	1.24

유효율의 변화량이 줄어 일정한 유효율로 수렴되는 경향을 보인다.

누수기록 및 유량자료의 패턴이 달라지면 유효율이 어떻게 변화하는지 분석하기 위해 시간의 역순으로 2012년 10월 자료부터 1개월씩 자료를 누적하면서 PCA를 실시하였다. Table 5는 최근 자료부터 2011년 9월까지 순차적으로 본 연구에서 개발한 알고리즘을 적용한 후 계산된 탐지 유효율을 나타낸다. Table 6은 Table 5에 보인 최고 유효율을 가지는 중심시간과 시간간격을 나타낸다.

Table 5를 분석하면 누적 3개월까지는 누수자료가 없는 블록이 많아 탐지유효율이 산정되지 않는 블록이 5개가 존재한다. 탐지유효율이 계산되는 블록들은 자료의 누적에 따라 유효율이 감소하는데, 이는 누수기록이 비교적 크게 증가하는 가을과 겨울을 제외한 다른 계절에서 발생하는 누수기록은 이상치가 발생한 날짜의 전후 7일 내에 들어오지 않기 때문인 것으로 보인다. 8개월 이후부터는 탐지유효율 100%에 가까워지고 있는데, 이는 누적 9개월부터 13개월(2011년 10월~2012년 2월)사이의 기간 동안 누수기록이 크게 증가하면서 탐지유효율이 상승하는 것으로 보인다.

모든 누수 및 유량자료가 포함된 14개월 자료의 경우 11개의 블록 중 6개의 블록에서 70%이상의 탐지유효율을 갖는데, 이러한 6개의 블록은 모두 9개 이상의 누수기록을 갖는 것으로 나타났다. 탐지유효율이 60% 이하로 낮게 나오는 F, G, H 블록의 경우 누수기록의 갯수가 다른 블록에 비해 그 수가 작은 편으로 나타

났다. 따라서 유량자료가 누적됨에 따라 안정적인 유효율을 얻기 위해서는 충분한 양의 누수기록이 필요함을 알 수 있다.

Table 4와 Table 6을 분석하면 중심시간은 월별 누수기록 개수가 전월에 비해 50%이상 증가하는 경우 비교적 큰 변화를 보인다. 또한 중심시간은 자료가 누적됨에 따라 자료누적 방법에 상관없이 19시에서 05시 사이의 해가 진 이후에 존재하였고, 시간간격은 K 블록을 제외한 다른 블록은 7시간 이내로 좁혀졌다.

유효율은 블록에 따라 많은 차이를 보이는데 그 중 K 블록과 같이 자료를 누적하는 방법에 관계없이 유효율이 100%를 초과하는 경우도 발생하였다. 이는 블록 내에 단기간에 연속적으로 기록된 누수기록이 존재하기 때문에 발생하는 것으로 사료된다.

4. 결 론

본 연구에서는 상수도 관망에서 관측된 유량자료를 이용하여 관망에서 발생하는 누수를 감지하기 위한 컴퓨터 계산 알고리즘을 개발하였다. 본 연구에서 개발된 알고리즘은 다변량 자료 통계분석 기법의 하나인 PCA에 기반한 것으로서, 기존의 24시간 관측된 유량자료 전체를 사용하는 기법(Palau et al., 2003)의 누수감지 효율을 향상시키기 위해 특정 시간대의 유량자료만을 사용해 PCA를 실시한 후 DMOD 통계치를 사용하여 산정되는 이상치와 누수기록 날짜와 비교하는 알고리즘이다.

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 009-016

pp. 159-167

Table 6. Center time and time duration that correspond to the maximum EODRs of Table 5

Block		month													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	Center Time	5	21	13	13	13	5	16	16	5	19	14	14	5	5
	Time Duration	3	3	9	9	5	3	7	7	3	3	3	3	3	3
	No. of Leaks	1	1	1	2	2	3	4	4	5	8	10	10	15	17
B	Center Time	5	19	19	20	20	20	5	20	7	7	17	20	12	12
	Time Duration	3	3	5	5	5	5	3	5	3	3	3	7	7	7
	No. of Leaks	2	2	2	3	3	4	4	4	4	7	9	10	11	12
C	Center Time	N/A	N/A	N/A	17	17	17	17	17	17	17	11	11	21	21
	Time Duration	N/A	N/A	N/A	17	17	17	17	17	17	17	5	5	7	7
	No. of Leaks	0	0	0	1	1	1	1	1	1	2	3	6	7	9
D	Center Time	1	1	12	11	12	12	12	12	12	12	12	12	4	6
	Time Duration	3	3	3	5	3	3	3	3	3	3	3	3	5	3
	No. of Leaks	1	1	1	1	1	1	1	1	1	1	2	2	5	5
E	Center Time	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	13	13	13	13	13	5
	Time Duration	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	3	3	3	3	3	3
	No. of Leaks	0	0	0	0	0	0	0	0	2	2	3	3	4	9
F	Center Time	N/A	N/A	N/A	24	24	24	22	22	23	23	21	21	24	24
	Time Duration	N/A	N/A	N/A	7	3	7	9	9	3	3	7	7	3	3
	No. of Leaks	0	0	0	1	1	2	5	5	5	5	6	6	6	6
G	Center Time	1	20	9	3	10	23	23	5	5	23	10	16	21	21
	Time Duration	3	11	3	5	3	5	5	3	3	5	3	3	3	3
	No. of Leaks	1	1	1	1	1	1	1	2	3	3	4	5	7	8
H	Center Time	N/A	N/A	N/A	1	24	23	23	24	24	24	1	20	20	16
	Time Duration	N/A	N/A	N/A	3	5	7	3	5	5	5	5	9	9	3
	No. of Leaks	0	0	0	1	1	1	1	1	1	1	2	2	2	2
I	Center Time	N/A	N/A	17	17	21	6	4	6	6	5	7	2	1	2
	Time Duration	N/A	N/A	3	3	13	5	5	5	3	5	3	3	3	3
	No. of Leaks	0	0	1	2	2	4	6	8	8	9	12	12	14	17
J	Center Time	N/A	N/A	N/A	N/A	N/A	12	12	3	12	7	7	10	1	1
	Time Duration	N/A	N/A	N/A	N/A	N/A	9	9	3	3	3	3	9	7	7
	No. of Leaks	0	0	0	0	0	1	1	1	1	2	3	5	8	10
K	Center Time	21	19	19	21	21	21	17	18	22	22	20	14	19	19
	Time Duration	3	11	7	7	7	7	5	13	3	3	3	5	13	13
	No. of Leaks	3	3	3	4	6	8	8	11	11	12	16	19	21	25

Table 3과 Table 5를 보면 PCA에 사용된 유량자료가 누적되면서 유효율의 변동이 존재하고 경우에 따라 누적된 유량자료의 개수가 적은 경우 오히려 유효율이 높은 결과를 보이기도 한다. 그러나 계산된 유효율의 신뢰성을 고려하면 PCA에 사용되는 유량자료의 개수가 많을수록 계산된 이상치 탐지 유효율의 신뢰성은 더 높을 것으로 사료된다.

본 연구의 결과 이상치 탐지 유효율이 가장 높게 계산되는 중심시간과 시간간격이 주로 19시에서 05시 사이의 야간 시간대로 구해진 것인데, 이러한 결과는 소블록 유입량 자료와 PCA를 이용한 누수감지 방법도 일반적으로 사용하는 야간 최소유량과 물수지 분석을 이용한 누수량 산정방법과 비슷하게 물 사용량이 적은 야간 시간대에 적용하는 것이 타당하다는 것을 나타낸다.



본 연구에서 개발된 기법은 상수도 관망에서 1년 이상 누적된 소블록 유입유량 자료와 누수기록을 이용하여 매일 본 연구에서 개발한 PCA알고리즘을 실시한 후 바로 전 날의 유량자료가 이상치로 계산되는 지를 확인하고, 이상치가 발견되는 경우 약 70%이상의 높은 이상치 탐지유효율을 가지는 블록에 한해 누수가 의심되는 구간에 대해 추가적인 누수검사를 실시하기 위한 의사결정에 도움을 줄 수 있을 것으로 사료된다. 본 연구에서 개발된 상수도 관망 누수감지 기법의 핵심은 컴퓨터 계산 알고리즘으로서 상수도 관망의 유량자료와 누수기록 자료가 모두 있다면 본 연구대상 지역 외 다른 상수도 관망의 누수감지에 적용이 가능할 것으로 사료된다. 다만, 각 상수도 블록의 특성에 따라 누수 감지의 정확도는 본 논문의 결과와는 상이할 수 있다.

이상치 탐지 유효율이 100%가 넘는 경우는 블록 내에서 누수기록이 약 7일 이내의 짧은 기간에 연속적으로 발생하여 하나의 이상치가 각 누수기록에 중복되어 인식되기 때문인 것으로 사료된다. 따라서 누수기록간의 간격이 짧은 경우 발생하는 탐지 범위의 중복을 막기 위해 이상치 탐지 알고리즘을 개선하는 연구가 필요한 것으로 사료된다.

사 사

본 연구는 환경부 “글로벌탑 환경기술개발사업 (2016002130005)”으로 지원 받은 과제입니다.

References

Borges, L.A. and Ramirez, M.A. (2010). “Acoustic Water Leak Detection System”, *Water Distribution Systems Analysis Symposium*, 1-3.

Covas, D., Ramos, H., Lopes, N. and Almeida, A.B. (2006). “Water pipe system diagnosis by transient pressure signals”, *Water Distribution Systems Analysis Symposium*, ASCE, USA, 1-19.

Kapelan, Z. Savic, D.A. and Walters, G.A. (2004). Incorporation of prior information on parameters in inverse transient analysis for leak detection and roughness calibration, *Urban*

Water J., 1(2), 129-143.

Mounce, S.R. Day, A.J. Wood, A.S. and Khan, A. (2002). A neural network approach to burst detection, *Water Sci. Technol.*, 45(4-5), 237-246.

Muggleton, J.M. and Brennan, M.J. (2005). Axisymmetric wave propagation in buried, fluid-filled pipes: effects of wall discontinuities, *J. Sound. Vib.*, 281(3-5), 849-867.

Muggleton, J.M. Brennan, M.J. Pinnington, R.J. and Gao, Y. (2006). A novel sensor for measuring the acoustic pressure in buried plastic water pipes, *J. Sound. Vib.*, 295(3-5), 1085-1098.

Palau, C.V. Arregui, F. and Ferrer, A. (2004). Using multivariate principal component analysis of injected water flows to detect anomalous behaviors in a water supply system. a case study, *Water Supply*, 4(3), 169-181.

Park, S. Jeon, D. Jung, S. Kim, J.H. and Lee, D. (2013). Identifying an appropriate analysis duration for the principal component analysis of water pipe flow data, *J. Korean Soc. Water Wastewater*, 27(3), 351-361.

Park, S. Jung, S.Y. and Sahleh, V. (2015). Determining the time of least water use for the major water usage types in district metered areas, *J. Korean Soc. Water Wastewater*, 29(3), 415-425.

Pilcher, R. Hamilton, S. Chapman, H. Ristovski, B. and Strapely, S. (2007). "Leak location and repair guidance notes", *International Water Association, Water Loss Task Forces: Specialist Group Efficient Operation and Management*, Bucharest, Romania, 12-18.

Stathis, J.A. and Loganathan, G.V. (1999). “Analysis of pressure-dependent leakage in water distribution systems, Analysis of pressure-dependent leakage in water distribution systems”, *Preparing for the 21st Century, 29th Annual Water Resources Planning and Management Conference*, Tempe, AZ, USA.

Tajima, M. and Mita, A. (2008). Automatic leakage detection for water supply systems using principal component analysis, *Materials Forum*, 33, 87-94.

Xia, L. and Guo-jin, L. (2010). “Leak detection of municipal water supply network based on the cluster-analysis and fuzzy pattern recognition”, *International Conference on E-Product, E-Service, and EEntertainment(ICEEE)*.

Ye, G and Fenner, Ra. (2011). Kalman filtering of hydraulic measurements for burst detection in water distribution systems, *J. Pipeline Syst. Eng. Pract.*, 2, 14-22.

pp. 009-016
pp. 009-016
pp. 009-016
pp. 009-016
pp. 009-016
pp. 009-016
pp. 009-016
pp. 009-016
pp. 159-167